泛函分析

下 册

〔苏〕Л. В. Канторович
Г. П. Акилов

高等教育出版社

泛函分析

下　册

[苏] Л.В.Канторович Г.П.Акилов
郭宜斌 译

高等教育出版社

内容提要

本书系根据Л. В. Канторович, Г. П. Акилов 著“泛函分析”第二版(1977 年)译出，分上、下两册出版.

本书为 1959 年出版的“赋范空间中的泛函分析”一书的修订本. 书中的叙述以一般泛函空间为基础，反映了这些年来在一系列问题上的进展. 这一版中泛函分析的应用在很大程度上得到了反映，除了在计算数学与数学物理上的应用外，还注意到对数理经济问题的某些应用.

上册为“泛函分析”的第一部分——线性算子与线性泛函，共十一章. 内容有拓扑空间与度量空间，向量空间，拓扑向量空间，赋范空间，线性算子与线性泛函，泛函的解析表示，线性算子序列，Banach 空间中的弱拓扑，紧算子与共轭算子，有序赋范空间，积分算子.

本书为“泛函分析”一书的第二部分——泛函方程，共七章. 主要内容为：共轭方程，第二类泛函方程，近似方法的一般理论，最速下降法，不动点原理，非线性算子的微分，Newton 法.

本书可作为高等学校数学专业、计算专业教师和研究生以及泛函专门组学生的教学参考书，也可供数学及其应用领域内的科学工作者使用.

泛函分析

下 册

[苏]Л. В. Канторович Г. П. Акилов

郭宜斌 译

*

高等教育出版社出版

新华书店北京发行所发行

四川新华印刷厂印装

*

开本 850×1168 1/32 印张 10.625 字数 257,000

1982 年 8 月第 1 版 1984 年 4 月第 1 次印刷

印数 00,001—10,800

书号 13010·0784 定价 1.65 元

目　　录

第二部分

泛函方程

第二部分

泛 函 方 程

第十二章 共轭方程

§1. 关于逆算子的定理*)

我们在本章中将对第一部分中（参见第五章§4）给出的关于逆算子的结论作一些补充讨论.

1.1. 回忆一下在第一部分中所给出的定义. 设U是从赋范空间$\mathbf{X}$到赋范空间$\mathbf{Y}$内的连续线性算子，如果存在从$\mathbf{Y}$到$\mathbf{X}$内的算子V，使得

$$VU=I_{\mathbf{X}} \quad (I_{\mathbf{X}}x=x, x\in\mathbf{X}), \tag{1}$$

$$UV=I_{\mathbf{Y}} \quad (I_{\mathbf{Y}}y=y, x\in\mathbf{Y}), \tag{2}$$

则称V是U的逆算子并写成$V=U^{-1}$.

逆算子U^{-1}的存在（即使不连续）等价于U实现从$\mathbf{X}$到$\mathbf{Y}$上的一对一的映射. 此外，如果U^{-1}是连续的，则此映射将是同构映射.

如果关系式(1)或(2)仅有一个成立，则称算子V是左逆算子或右逆算子并分别记为$V=U_l^{-1}$，$V=U_r^{-1}$**). 在V.4.4中已经证明：连续左逆算子存在的充分必要条件是

$$\|U(x)\|\geqslant m\|x\| \qquad (x\in\mathbf{X}), \tag{3}$$

其中m是不依赖于x的正数. 并且，如果算子U映$\mathbf{X}$到全空间$\mathbf{Y}$上，则左、右逆算子均存在，即此时存在连续（双边）算子U^{-1}.

1.2. 我们来证明下述定理.

---

*) 本节的内容参见 Neumann[1]或 Schauder[2].

**) 有时符号l和r将略去不写.

定理1. 如果将 B-空间 $\mathbf{X}$ 变换到赋范空间 $\mathbf{Y}$ 内的连续线性算子 U 有连续左逆算子，则集合 $\mathbf{Y}'=U(\mathbf{X})$ 是 B-空间.

证. 这里只要证明空间 $\mathbf{Y}'$ 的完备性. 设 $\{y_n\}$ 是 $\mathbf{Y}'$ 中的自收敛序列. 令 $x_n=U^{-1}(y_n)$ $(y_n=U(x_n); n=1,2,\cdots)$. 据 1.1 中所述可知

$$\|y_n-y_k\|=\|U(x_n)-U(x_k)\|\geqslant m\|x_n-x_k\|,$$

其中 m 是正常数. 于是

$$\lim_{k,n\to\infty}\|x_n-x_k\|=0.$$

因而在 $\mathbf{X}$ 中存在元素 $x_0=\lim\limits_{n\to\infty}x_n$. 因为 $\lim y_n=\lim U(x_n)=U(x_0)$，命 $y_0=U(x_0)$，则得 $y_0\in\mathbf{Y}'$ 及 $y_n\to y_0$. 空间 $\mathbf{Y}'$ 的完备性于是得证.

推论. 在上述定理的条件下，$\mathbf{Y}'$ 是 $\mathbf{Y}$ 中的闭集合.

1.3. 设 U 是从赋范空间 $\mathbf{X}$ 到赋范空间 $\mathbf{Y}$ 内的连续线性算子. 显而易见，集合 $\mathbf{X}_0=U^{-1}(\mathbf{0})$ 是空间 $\mathbf{X}$ 的闭子空间. 由 $\mathbf{X}$ 和 $\mathbf{X}_0$ 构成商空间 $\overline{\mathbf{X}}=\mathbf{X}/\mathbf{X}_0$（参见 IV.1.8）. 设 $\bar{x}\in\overline{\mathbf{X}}$；考虑任意元素 $x\in\bar{\mathbf{X}}$，令

$$\bar{U}(\bar{x})=U(x). \tag{4}$$

因若 $x',x''\in\bar{x}$，则 $x'-x''\in\mathbf{X}_0$，因而 $U(x')=U(x'')$. 可见元素 $\bar{U}(\bar{x})$ 的定义和元素 $x\in\bar{x}$ 的选择无关. 这样，式(4)就定义了从空间 $\mathbf{X}$ 到 $\mathbf{Y}$ 内的算子 $\bar{U}$. 这个算子是齐次的和可加的. 它还是连续的：对不等式

$$\|\bar{U}(\bar{x})\|=\|U(x)\|\leqslant\|U\|\|x\| \qquad (x\in\bar{x}),$$

的右端取下确界可得

$$\|\bar{U}(\bar{x})\|=\|U(x)\|\leqslant\|U\|\|\bar{x}\| \qquad (\bar{x}\in\overline{\mathbf{X}}).$$

由算子 $\bar{U}$ 的定义本身推知 $U=\bar{U}\varphi$，其中 φ 是从 $\mathbf{X}$ 到 $\mathbf{X}/\mathbf{X}_0$ 的自然同态（见 IV.1.8）而且 $\|U\|=\|\bar{U}\|$.

算子 $\bar{U}$ 和 U 的不同在于它实现从空间 $\overline{\mathbf{X}}$ 到空间 $\mathbf{Y}$ 内的一对一

的对应(映射). 事实上, 如果 $\bar{U}(\bar{x})=0$, 则对于任意的 $x\in\bar{x}$ 将有 $U(x)=0$, 亦即 $x\in U^{-1}(0)=\mathbf{X}_0$, 因而 $\bar{x}$ 和空间 $\overline{\mathbf{X}}$ 中的零元素组成的类 $\mathbf{X}_0$ 一致.

当 U 将 $\mathbf{X}$ 映射到 $\mathbf{Y}$ 上时, 算子 $\bar{U}$ 将 $\overline{\mathbf{X}}$ 映射到 $\mathbf{Y}$ 上. 此外, 如果存在连续的逆算子 $\bar{U}^{-1}$, 则称空间 $\mathbf{X}$ 和空间 $\mathbf{Y}$ 是同态的, 算子 U 称作是从空间 $\mathbf{X}$ 到空间 $\mathbf{Y}$ 上的同态算子.

用原始空间 $\mathbf{X}$ 和 $\mathbf{Y}$ 的术语来讲, 从空间 $\mathbf{X}$ 到空间 $\mathbf{Y}$ 上的同态算子 U 可由下述两个条件来表征:

1) $U(\mathbf{X})=\mathbf{Y}$;

2) 存在 $m>0$, 对于每一个 $y\in\mathbf{Y}$, 可求得 $x\in\mathbf{X}$ 使得

$$y=U(x),\ \|y\|\geqslant m\|x\|.$$

事实上, 如果 U 是同态算子, 则1)显然成立. 其次, 由 IV.1.8 的式(4), 取 $x\in\bar{x}=U^{-1}(y)$, 则得

$$\|x\|\leqslant 2\|\bar{x}\|\leqslant 2\|\bar{U}^{-1}\|\|y\|,$$

因而可取

$$m=\frac{1}{2\|\bar{U}^{-1}\|}.$$

反之, 如果两个条件均满足, 则从 1) 得 $\bar{U}(\mathbf{X})=\mathbf{Y}$. 设 $\bar{x}\in\overline{\mathbf{X}}$, 对 $y=\bar{U}(\bar{x})$ 求与第二个条件相应的 $x\in\mathbf{X}$. 由于

$$\bar{U}(\varphi(x))=U(x)=y=\bar{U}(\bar{x}),$$

以及算子 $\bar{U}$ 的相互单值性, 可知 $\varphi(x)=\bar{x}$. 因此

$$\|\bar{U}(\bar{x})\|=\|y\|\geqslant m\|x\|\geqslant m\|\bar{x}\|.$$

正如在 1.1 中所指出的那样, 这个不等式及关系 $\bar{U}(\overline{\mathbf{X}})=\mathbf{Y}$ 就保证存在连续(双边)逆算子 $\bar{U}^{-1}$.

1.4. 定理 1 的逆命题在泛函方程理论中起着很基本的作用, 它是下述引理的一部分.

引理 1. 设 $\bar{U}$ 是从 B-空间 $\mathbf{X}$ 到赋范空间 $\mathbf{Y}$ 内的连续线性算

子．这时，假设B表示空间$\mathbf{X}$的以零点为中心的单位球，以S_r表示空间$\mathbf{Y}$中以零点为中心以$r>0$为半径的球．如果象$U(B)$在S_r中稠密，则U是从空间$\mathbf{X}$到空间$\mathbf{Y}$的同态算子．特别，如果算子U所实现的映射是一对一的话，则存在连续的双边逆算子U^{-1}.

证．我们来验证上一段中的二个条件是满足的．显然，可以认为球B和S_r均是闭球；今证

$$U(B)\supset S_{r/2}. \tag{5}$$

取正数序列$\{\varepsilon_n\}$，使得$\sum_{k=1}^{\infty}\varepsilon_k\leqslant 1$，并考察$y\in S_r$．因为$\overline{U(B)}\supset S_r$，则可找到$y_1\in U(B)$，使得

$$\|y-y_1\|\leqslant\varepsilon_1 r.$$

用x_1表示B中使得$y_1=U(x_1)$的元素．设B_h是空间$\mathbf{X}$中以零点为中心以h为半径的闭球．由引理的条件推知$\overline{U(B_h)}\supset S_{hr}$．因此，由于元素$y-y_1\in S_{\varepsilon_1 r}$，故可找到元素$x_2\in B_{\varepsilon_1}$使得

$$\|y-(y_1+y_2)\|\leqslant\varepsilon_2 r \qquad (y_2=U(x_2)).$$

用类似的办法重复以上的论证，我们就得到满足如下条件的两个序列$\{y_n\}\subset\mathbf{Y}$和$\{x_n\}\subset\mathbf{X}$:

$$y_n=U(x_n),\ x_n\in B_{\varepsilon_{n-1}},\ \left\|y-\sum_{k=1}^{n}y_k\right\|\leqslant\varepsilon_n r \tag{6}$$

$$(n\in\boldsymbol{N},\ \varepsilon_0=1).$$

由于$\|x_n\|\leqslant\varepsilon_{n-1}$而$\mathbf{X}$是完备空间，所以级数$\sum_{k=1}^{\infty}x_k$收敛．如果用$x$表示这个级数的和，则

$$\|x\|\leqslant\sum_{k=1}^{\infty}\|x_k\|\leqslant\sum_{k=1}^{\infty}\varepsilon_{k-1}\leqslant 2,$$

也就是说$x\in B_2$．其次，

$$U(x)=\sum_{k=1}^{\infty}U(x_k)=\sum_{k=1}^{\infty}y_k.$$

但从式(6)显然可见$\sum_{k=1}^{\infty}y_k=y$. 因此 $y=U(x)$. 这样,我们就证明了 $U(B_2)\supset S_r$,这与式(5)等价.

因由(5)推知 $U(B_n)\supset S_{nr/2}$,所以

$$U(\mathbf{X})=\bigcup_{n=1}^{\infty}U(B_n)\supset\bigcup_{n=1}^{\infty}S_{nr/2}=\mathbf{Y},$$

这就验证了第一个条件.

其次,如果 $y\neq\mathbf{0}$ 是 $\mathbf{Y}$ 中任意一个元素,则

$$y'=\frac{r}{2\|y\|}y\in S_{r/2},$$

并且由(5)可以找到元素 $x'\in B$ 使得 $y'=U(x')$. 命 $x=\frac{2\|y\|}{r}x'$, 则得到

$$U(x)=y,\ \|x\|=\frac{2}{r}\|y\|\|x'\|\leqslant\frac{2}{r}\|y\|.$$

因此,第二个条件亦满足. 引理得证.

推论. 如果引理的条件满足,则 $\mathbf{Y}$ 是完备空间.

事实上,由于 U 是同态算子, 故从商空间 $\overline{\mathbf{X}}=\mathbf{X}/\mathbf{X}_0$ $(\mathbf{X}_0=U^{-1}(0))$ 到 $\mathbf{Y}$ 的算子 $\bar{U}$ 有连续的逆算子. 因为 $\mathbf{Y}=\bar{U}(\overline{\mathbf{X}})$, 应用定理 1 立得此推论.

引理的条件验证起来比较困难, 更为方便的是下述定理的条件.

定理 2(Banach). 如果$U(\mathbf{X})$是空间 $\mathbf{Y}$ 中的第二纲集合, 则引理的条件得以满足, 从而U 是从空间 $\mathbf{X}$ 到空间 $\mathbf{Y}$ 上的同态算子.

证. 仍用引理中各记号, 我们来证明: 如果引理的条件不满

足，则集合 $U(B)$ 处处不稠密．设若不然，则可求得空间 $\mathbf{Y}$ 中的以点 $y_0\in\mathbf{Y}$ 为中心以 r 为半径的球 $S(y_0, r)$ 使得

$$\overline{U(B)}\supset S(y_0, r). \tag{7}$$

集合 $U(B)$ 是对称的，即它同时包含元素 y 和 $(-y)$．显然，闭包 $\overline{U(B)}$ 同样是对称的．于是由(7)式可得

$$\overline{U(B)}\supset S(-y_0, r).$$

取 $y\in S_r$．元素 (y_0+y) 和 $(-y_0+y)$ 分别属于 $S(y_0, r)$ 和 $S(-y_0, r)$，因之这两个元素也都属于 $\overline{U(B)}$．但 $U(B)$ 是凸集，因此 $\overline{U(B)}$ 也是凸集，故 $\overline{U(B)}$ 应包含分别属于这两个集合的元素的半和．特别是有

$$y=\frac{1}{2}((y_0+y)+(-y_0+y))\in\overline{U(B)}.$$

于是 $\overline{U(B)}\supset S_r$．

于是，如果引理 1 的条件不满足，集合 $U(B)$ 处处不稠密．同样可证对 $U(B_n)$ $(n\in N)$ 中任何一个也都如此．但因为

$$U(\mathbf{X})=\bigcup_{n=1}^{\infty}U(B_n),$$

故得知 $U(\mathbf{X})$ 是第一纲集合．

定理得证．

我们再指出这个定理的下述推论，它是定理 1 的逆定理．

推论． 如果连续线性算子 U 实现了从 B-空间 $\mathbf{X}$ 到 B-空间 $\mathbf{Y}$ 的闭子空间上的一对一的映射，则逆算子 U^{-1} 是连续算子．

事实上，B-空间的闭子空间仍然是 B-空间，因而它本身是第二纲集(见 I. 4. 7)．

1. 5. 下面我们指出定理 2 的某些直接应用．

假定在向量空间 $\mathbf{X}$ 中用两种不同的办法引进范数．分别以 $\|x\|_1$ 和 $\|x\|_2$ 表示在第一种和第二种范数定义方法之下元素 $x\in\mathbf{X}$

的范数. 因此, 集合 $\mathbf{X}$ 就变成两个赋范空间 $\mathbf{X}_1$ 和 $\mathbf{X}_2$. 尽管 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 应看成两个不同的赋范空间, 但它们之间可能不存在实质上的区别. 例如当在一个空间中收敛的序列在另一个空间中也收敛于同一个元素时就是如此. 这时, 我们就说空间 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 中的范数是等价的, 这表明空间 $\mathbf{X}_1$ 和空间 $\mathbf{X}_2$ 是同构的(参见 IV.1.3).

定理 3. 设 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 是两个 B-空间, 并且 $\mathbf{X}_1 \subset \mathbf{X}_2$ *). 如果从在空间 $\mathbf{X}_1$ 中 $x_n \to x$ 可推知在空间 $\mathbf{X}_2$ 中 $x_n \to x$, 则或者 $\mathbf{X}_1 = \mathbf{X}_2$ 且空间 $\mathbf{X}_1$ 和 $\mathbf{X}_2$ 中的范数等价, 或者 $\mathbf{X}_1$ 是 $\mathbf{X}_2$ 中的第一纲集合.

证. 用 U 表示从空间 $\mathbf{X}_1$ 到空间 $\mathbf{X}_2$ 内的嵌入算子, 也就是说 U 是这样的算子: 它将元素 $x \in \mathbf{X}_1$ 映射为同一个 x, 但看成是 $\mathbf{X}_2$ 中的元素. 易见 U 是连续线性算子. 如果集合 $\mathbf{X}_1 = U(\mathbf{X}_1)$ 是空间 $\mathbf{X}_2$ 中的第二纲集, 则由定理 2 知总有 $\mathbf{X}_1 = \mathbf{X}_2$ 并且 U 有连续的逆算子. 因而, 若在空间 $\mathbf{X}_2$ 中有极限关系 $x_n \to x$ 存在, 则在空间 $\mathbf{X}_1$ 中存在这样的极限关系: $x_n = U^{-1}(x_n) \to U^{-1}(x) = x$. 这就是说空间 $\mathbf{X}_1$ 和空间 $\mathbf{X}_2$ 中的范数是等价的.

注. 象上面一样, 分别用 $\|x\|_1$ 和 $\|x\|_2$ 表示元素 x 在空间 $\mathbf{X}_1$ 中和在空间 $\mathbf{X}_2$ 中的范数, 则范数等价这个事实可以叙述成: 存在常数 $m, M > 0$ 使得

$$m\|x\|_1 \leqslant \|x\|_2 \leqslant M\|x\|_1 \qquad (x \in \mathbf{X}_1 = \mathbf{X}_2).$$

事实上, 命 $M = \|U\|$ 和 $m = \dfrac{1}{\|U^{-1}\|}$ 即可.

将定理应用到 $\mathbf{X}_1 = \mathbf{C}^{(1)}(D)$ 和 $\mathbf{X}_2 = \mathbf{C}(D)$ 的情形, 我们就得到下述结论: 所有连续可微函数的集合是连续函数空间 $\mathbf{C}$ 中的第一纲集. 同样可知: 几乎处处有界可测函数的集合是空间 $\mathbf{L}^1$ 中的第一纲集, 等等.

---

*) 这时我们假定嵌入关系 $\mathbf{X}_1 \subset \mathbf{X}_2$ 保持代数运算, 即 $\mathbf{X}_1$ 可看作是 $\mathbf{X}_2$ 中的线性集.

1.6. 设T是从赋范空间$\mathbf{X}$的集合Ω到赋范空间$\mathbf{Y}$上的算子(不一定是线性的),如果从

$$x_n\in\Omega \quad (n=1,2,\cdots); \qquad x_n\to x_0,\ T(x_n)\to y_0,$$

可推出$x_0\in\Omega$且$T(x_0)=y_0$,则称T是闭算子.

定义于闭集合上的连续线性算子显然是闭算子. 若T是线性闭算子而$\mathbf{X}$和$\mathbf{Y}$均为B-空间,则逆命题亦成立,即下述定理成立.

定理 4. 设T是从B-空间$\mathbf{X}$的线性闭集合Ω到B-空间$\mathbf{Y}$内的线性闭算子,则T是连续算子.

证. 在定理的假设之下因Ω本身也是B-空间,故可认为$\Omega=\mathbf{X}$. 在空间$\mathbf{X}$中引进新的范数,命

$$\|x\|_1=\|x\|+\|T(x)\| \qquad (x\in\mathbf{X}). \tag{8}$$

不难验证这样的范数定义满足赋范空间的公理. 今证空间$\mathbf{X}$对于新范数是完备的. 设

$$\lim_{k,n\to\infty}\|x_n-x_k\|_1=0.$$

这表明

$$\lim_{k,n\to\infty}\|x_n-x_k\|=0 \quad 而且 \quad \lim_{k,n\to\infty}\|T(x_n)-T(x_k)\|=0.$$

因为对这样的范数而言空间$\mathbf{X}$和$\mathbf{Y}$都是完备的,所以必存在极限

$$\lim_{n\to\infty}x_n=x \quad 及 \quad \lim_{n\to\infty}T(x_n)=y_0.$$

由算子T的闭性,得知$y_0=T(x_0)$. 但这时

$$\lim_{n\to\infty}\|x_n-x_0\|_1=\lim_{n\to\infty}\|x_n-x_0\|+\lim_{n\to\infty}\|T(x_n)-T(x_0)\|=0$$

这就证明了空间$\mathbf{X}$关于新范数是完备的.

因为

$$\|x\|\leqslant\|x\|_1,$$

所以由$\|x_n\|_1\to 0$可推知$\|x_n\|\to 0$. 利用上一定理,我们就得到

$$\|x\|_1\leqslant M\|x\|.$$

更有

$$\|Tx\| \leqslant M\|x\|,$$

这表明算子T是连续的.

注. 定义于整个空间上或闭线性集合上的闭线性算子类与连续线性算子类是重合的. 但若考察在线性(非闭)集上的闭线性算子,则它们实际上构成了较连续算子类更广的一类算子. 例如空间$\mathbf{L}^2(a,b)$中的算子T:

$$y=T(x),\ \ y(t)=\frac{dx(t)}{dt},$$

是定义在所有其导数属于$\mathbf{L}^2(a,\ b)$的绝对连续函数的集合Ω上的,不难验证,T是闭算子但不是连续的.

闭的但非连续的算子主要是在$\mathbf{X}=\mathbf{Y}$是Hilbert空间时加以研究. 一般情形的研究十分困难,因为这要涉及到任意B-空间的结构问题.

§2. 已给方程与其共轭方程之间的联系

本节我们同时考察两个有关联的方程

$$U(x)=y, \tag{1}$$

$$U^*(g)=f. \tag{2}$$

我们称方程(2)是方程(1)的共轭方程. 通常假定U是从空间$\mathbf{X}$到空间$\mathbf{Y}$内的连续线性算子. 尽管本节的某些定理不要求$\mathbf{X}$和$\mathbf{Y}$一定是完备空间,但我们以后总认为$\mathbf{X}$和$\mathbf{Y}$是完备的.

本节的定理对于空间$\mathbf{L}^2(a,\ b)$的情形是由Hellinger和Toeplitz [1],[2]证明的;对于空间$\mathbf{l}^p$和$\mathbf{L}^p(a,b)$的情形是由Riesz[3]证明的;一般情形的证明可参见Banach;Zaanen-I.

2.1. 在下面的叙述中,所谓零集合将起很大的作用. 设Γ是空间$\mathbf{X}$中线性泛函的某个集合. 用$\mathbf{N}(\Gamma)$表示对任何$f\in\Gamma$使得

$f(x)=0$ 的所有 $x\in\mathbf{X}$ 的集合. 如果 $E\subset\mathbf{X}$,则用 $\mathbf{N}^*(E)$ 表示所有这样的泛函的集合,它在集 E 中每个元素上取值为零. 集合 $\mathbf{N}(\Gamma)$ 和 $\mathbf{N}^*(E)$ 分别叫做集合 Γ 和集合 E 的零集合.

如果我们考察对偶对 $\langle\mathbf{X},\mathbf{X}^*\rangle$, 则显然 $\mathbf{N}(\Gamma)$ 是在 III. 3.2 中作为一般情况而引进的集合 Γ 的零化子 $\Gamma^\perp$,而 $\mathbf{N}^*(E)$ 是集合 E 的零化子 $E^\perp$. 根据 III. 3.2 中所叙述的零化子和极的性质, 以及在空间 $\mathbf{X}$ 中关于范数闭和弱闭的一致性, 我们就得到: 1) $\mathbf{N}(\Gamma)$ 是 $\mathbf{X}$ 中的闭线性集, 2) $\mathbf{N}(\mathbf{N}^*(E))$ 是 E 的闭线性包, 3) 如果 $\mathbf{X}_0$ 是 $\mathbf{X}$ 中的闭子空间,则

$$\mathbf{X}_0=\mathbf{N}(\mathbf{N}^*(\mathbf{X}_0)).$$

类似地可以得到: 1) $\mathbf{N}^*(E)$ 是 $\mathbf{X}^*$ 中的 $(*)$-弱闭线性子集, 2) $\mathbf{N}^*(\mathbf{N}(\Gamma))$ 是 Γ 的 $(*)$-弱闭线性包, 3) 如果 $\mathbf{Z}$ 是 $\mathbf{X}^*$ 中的 $(*)$-弱闭子空间,则

$$\mathbf{Z}=\mathbf{N}^*(\mathbf{N}(\mathbf{Z})).$$

因为, 若 $\mathbf{X}$ 是非自反的, 则在 $\mathbf{X}^*$ 中不是所有的关于范数闭的子空间是 $(*)$-弱闭的, 所以不是所有的闭子空间均具有形式 $\mathbf{N}^*(E)\,(E\subset\mathbf{X})$.

设 π: $\mathbf{X}\to\mathbf{X}^{**}$ 是典型嵌入. 如果 E 是空间 $\mathbf{X}$ 中的某个集合,则用 $E_0{}^{**}$ 表示 $\pi(E)$,即 $E_0{}^{**}$ 是 $\mathbf{X}^*$ 中所有具有形式 $F_x\,(x\in\mathbf{X})$ 的泛函集合,其中象通常那样

$$F_x(f)=\overline{f(x)}\qquad(f\in\mathbf{X}^*).$$

对于空间 $\mathbf{Y}$ 用类似的记号 $G_y\,(y\in\mathbf{Y})$.

显然,下式成立:

$$\mathbf{N}^*(E)=\mathbf{N}(E_0{}^{**}).$$

对于集合 $\Gamma\subset\mathbf{X}^*$, 可以考察两个零集合——$\mathbf{N}(\Gamma)\subset\mathbf{X}$ 和 $\mathbf{N}^*(\Gamma)\subset\mathbf{X}^{**}$. 容易看出

$$[\mathbf{N}(\Gamma)]_0{}^{**}=\mathbf{N}^*(\Gamma)\cap\mathbf{X}_0{}^{**},$$

其中 $\mathbf{X}_0^{**}$ 对应于通常记号，它表示集合 $\pi(\mathbf{X})$.

最后引进算子 U 的零集合 $\mathbf{N}(U)$，它是在算子 U 作用下变为零向量的所有 $x\in\mathbf{X}$ 的全体，换言之

$$\mathbf{N}(U)=U^{-1}(\mathbf{0}).$$

2.2. 我们要问：空间 $\mathbf{X}$ 的象 $U(\mathbf{X})$ 是由怎样的向量(点)构成的？下述定理部分地回答了这个问题.

定理1. 记 $\mathbf{Y}'=U(\mathbf{X})$ 并设 $\mathbf{Y}_1$ 是集合 $\mathbf{Y}'$ 的闭包. 那么

$$\mathbf{Y}_1=\mathbf{N}(\mathbf{N}(U^*)),$$

也就是说 $\mathbf{Y}_1$ 是使得算子 U^* 变为零的泛函 $g\in\mathbf{Y}^*$ 的公共零元素集合.

证. 首先证明关系式

$$\mathbf{N}^*(\mathbf{Y}')=\mathbf{N}(U^*). \tag{3}$$

设 $g\in\mathbf{N}^*(\mathbf{Y}')$. 因对任何 $x\in\mathbf{X}$ 有 $U(x)\in\mathbf{Y}'$，所以

$$U^*(g)(x)=g(U(x))=0 \qquad (x\in\mathbf{X}), \tag{4}$$

即 $U^*(g)=\mathbf{0}$ 且 $g\in\mathbf{N}(U^*)$. 等式(4)还表明：如果 $U^*(g)=\mathbf{0}$，则 $g\in\mathbf{N}^*(\mathbf{Y}')$. 这样，我们就证明了关系式(3)的成立. 对(3)式两端取其零集(即以算子 $\mathbf{N}$ 作用两端)后得到

$$\mathbf{N}(\mathbf{N}^*(\mathbf{Y}'))=\mathbf{N}(\mathbf{N}(U^*)).$$

但因 $\mathbf{Y}'$ 是线性集合，正如 2.1 中所指出的，其闭包与 $\mathbf{N}(\mathbf{N}^*(\mathbf{Y}'))$ 重合. 定理于是得证.

对偶的定理也成立.

定理 1*. 用 $\tilde{\mathbf{X}}^*$ 表示集合 $U^*(\mathbf{Y}^*)$ 的 (*)-弱闭包，则有

$$\tilde{\mathbf{X}}^*=\mathbf{N}^*(\mathbf{N}(U)).$$

证. 在(3)式中以 U^* 代替 U 后得到

$$\mathbf{N}^*(U^*(\mathbf{Y}^*))=\mathbf{N}(U^{**}),$$

并考察此式两端与集合 $\mathbf{X}_0^{**}$ 的交. 因为

$$U^{**}(F_x)(g)=F_x(U^*(g))=g(U(x))=G_{U(x)}(g)$$

$$(g\in\mathbf{Y}^*,\ x\in\mathbf{X}),$$

亦即

$$U^{**}(F_x)=G_{U(x)}\qquad(x\in\mathbf{X}),$$

交集 $\mathbf{N}(U^{**})\cap\mathbf{X}_0^{**}$ 由使得 $x\in\mathbf{N}(U)$ 的泛函 $F_x\in\mathbf{X}_0^{**}$ 所组成. 换言之,

$$\mathbf{N}(U^{**})\cap\mathbf{X}_0^{**}=[\mathbf{N}(U(\mathbf{Y}^*))]_0^{**}.$$

另一方面又有

$$\mathbf{N}^*(U^*(\mathbf{Y}^*))\cap\mathbf{X}_0^{**}=[\mathbf{N}(U^*(\mathbf{Y}^*))]_0^{**}.$$

比较所得到的关系式可得

$$\mathbf{N}(U^*(\mathbf{Y}^*))=\mathbf{N}(U),$$

因而

$$\mathbf{N}^*(\mathbf{N}(U^*(\mathbf{Y}^*)))=\mathbf{N}^*(\mathbf{N}(U)).$$

但是 $\tilde{\mathbf{X}}^*=\mathbf{N}^*(\mathbf{N}(U^*(\mathbf{Y}^*)))$, 于是定理得证.

2.3. 下面两个定理给出了方程(1)或(2)中一个方程对于任意右端项的可解性与另一个方程的齐次可解性之间的相依关系.

定理 2. 方程(2)对任意的$f\in\mathbf{X}^*$有解的充分必要条件是算子U具有连续的左逆算子U^{-1}.

证. 必要性. 设方程(2)对任意的 $f\in\mathbf{X}^*$ 均有解, 换言之, 设 $U^*(\mathbf{Y}^*)=\mathbf{X}^*$. 考察空间$\mathbf{X}$中任意非零元素 x. 由定理 V.7.2, 存在泛函 $f\in\mathbf{X}^*$使得

$$f(x)=\|x\|,\ \|f\|=1.$$

其次, 由定理 1.2, U^* 是同态算子, 因此可求得 $g\in\mathbf{Y}^*$使得

$$U^*(g)=f,\qquad\|g\|\leqslant m\|f\|=m,$$

其中m不依赖于 f, 因而也不依赖于 x. 综上可得

$$\|x\|=f(x)=g(U(x))\leqslant\|g\|\|U(x)\|\leqslant m\|U(x)\|.$$

因此,

$$\|U(x)\| \geqslant \frac{1}{m}\|x\| \qquad (x \in \mathbf{X}),$$

由此推知存在连续左逆算子 U^{-1}.

充分性. 设 f_0 是 $\mathbf{X}$ 中的任意线性泛函, 即 $f_0 \in \mathbf{X}^*$. 假定存在连续算子 $U_l^{-1} = U^{-1}$ 并记

$$g'(y) = f_0(U^{-1}(y)) \qquad (y \in U(\mathbf{X})).$$

定义于 $\mathbf{Y}' = U(\mathbf{X})$ 上的泛函 g' 是连续的和线性的. 将 g' 扩张到全 $\mathbf{Y}$ 上得到泛函 $g_0 \in \mathbf{Y}^*$. 因为对任意的 $x \in \mathbf{X}$ 均有 $U(x) \in \mathbf{Y}'$, 所以

$$U^*(g_0)(x) = g_0(U(x)) = g'(U(x)) = f_0(U^{-1}U(x)) = f_0(x).$$

即 $U^*(g_0) = f_0$, 亦即方程(2)对任意右端都可解.

定理 2* 方程(1)对任何 $y \in \mathbf{Y}$ 有解的充分必要条件是算子 U^* 具有连续左逆算子 U^{*-1}.

证. 必要性. 本定理所给条件的必要性的验证和定理 2 中条件的必要性验证类似. 设 $y \in \mathbf{Y}$ 是使得 $\|y\| \leqslant 1$ 的任意元素. 由定理 1.2, 可找到 $x \in \mathbf{X}$ 使得

$$U(x) = y, \ \|x\| \leqslant m\|y\| \leqslant m.$$

其次, 如果 $g \in \mathbf{Y}^*$ 且 $f = U^*(g)$, 则

$$|g(y)| = |g(U(x))| = |f(x)| \leqslant \|f\| \|x\| \leqslant m\|f\|.$$

在上式左端关于 $y(\|y\| \leqslant 1)$ 取上确界并注意到 m 不依赖于 y, 乃得

$$\|g\| = \sup_{\|y\| \leqslant 1} |g(y)| \leqslant m\|f\| = m\|U^*(g)\|,$$

这就得到了保证存在连续左逆算子 U^{*-1} 的条件.

充分性. 我们来证明算子 U 满足上节中引理 1 的条件, 因而由引理推知 $U(\mathbf{X}) = \mathbf{Y}$.

于是, 设 B 是空间 $\mathbf{X}$ 中以零点为中心的单位球. 集合 $U(B)$ 是绝对凸的, 因此它的闭包(在空间 $\mathbf{Y}$ 中的闭包)是 $U(B)^{\circ\circ}$(参见 III. 3.2). 极 $U(B)^{\circ}$ 是所有这样的泛函 g 的集合: $g \in \mathbf{Y}^*$, 对所有的

$y\in U(B)$，$|g(y)|\leqslant 1$．换句话说是这样的泛函 g 的集合：对所有的 $x\in B$，有 $|g(U(x))|\leqslant 1$．记 $f=U^*(g)$，则上面的后一个不等式可改写成

$$|f(x)|\leqslant 1 \qquad (x\in B).$$

这样，当 g 遍历集合 $U(B)^\circ$时，泛函 $f=U^*(g)$不会超出空间 $\mathbf{X}^*$的单位球 B°的范围，亦即

$$U^*(U(B)^\circ)\subset B^\circ\cap U^*(\mathbf{Y}^*).$$

将上面的关系式改写成如下的形式

$$U(B)^\circ\subset U^{*-1}(B^\circ\cap U^*(\mathbf{Y}^*)).$$

由于算子 U^{*-1} 的连续性，集合 $S=U^{*-1}(B^\circ\cap U^*(\mathbf{Y}^*))$在空间 $\mathbf{Y}^*$ 中有界．因此，极 S°就是空间$\mathbf{Y}$的零邻域．但是

$$\overline{U(B)}=(U(B))^{\circ\circ}\supset S^\circ,$$

这正是所要证明的．

从定理 2 和定理 2* 推出下面的推论．

推论．连续双边逆算子 U^{-1} 存在的充分必要条件是连续双边逆算子 U^{*-1} 存在，这时

$$U^{*-1}=(U^{-1})^*.$$

今给出上面等式的证明．设 $f\in\mathbf{X}^*$，$g=(U^{-1})^*(f)$，则

$$g(y)=f(U^{-1}(y))=f(x) \quad (y=U(x)\in\mathbf{Y}).$$

另一方面，若写 $g'=U^{*-1}(f)$，即 $f=U^*(g')$，则

$$f(x)=g'(U(x))=g'(y) \quad (x=U^{-1}(y)\in\mathbf{X}).$$

于是，对任何 $y\in\mathbf{Y}$ 有 $g(y)=g'(y)$，换言之，即有

$$(U^{-1})^*(f)=U^{*-1}(f) \quad (f\in\mathbf{X}^*),$$

这正是所要证明的．

2.4. 下面两个定理建立了方程(1)和方程(2)可解性的条件，这两个定理是定理 2 和定理 2* 的推广．

定理 3．如果 $U^*(\mathbf{Y}^*)$ 是闭集，则

$$U(\mathbf{X})=\mathbf{N}(\mathbf{N}(U^*)),$$

即方程(1)可解的充分必要条件是从 $U^*(y)=\mathbf{0}$ 推出 $g(y)=0$.

证. 考察空间 $\mathbf{Y}$ 的子空间 $\mathbf{Y}_1=\overline{U(\mathbf{X})}$ 以及从空间 $\mathbf{X}$ 到 $\mathbf{Y}_1$ 内的算子 U_1:

$$U_1(x)=U(x) \quad (x\in\mathbf{X})^{*)}.$$

我们来证明

$$U_1^*(\mathbf{Y}_1^*)=U^*(\mathbf{Y}^*). \tag{5}$$

为此,用 ω 表示空间 $\mathbf{Y}_1$ 嵌入于空间 $\mathbf{Y}$ 的嵌入算子. 显然,由于存在连续左逆算子 ω^{-1}, 根据定理 2 可知有 $\omega^*(\mathbf{Y}^*)=\mathbf{Y}_1^*$. 但 $U=\omega U_1$,所以 $U^*=U_1^*\omega^*$(参见 IX. 3. 1),因而

$$U^*(\mathbf{Y}^*)=U_1^*(\omega^*(\mathbf{Y}^*))=U_1^*(\mathbf{Y}_1^*).$$

因为 $\overline{U_1(\mathbf{X})}=\mathbf{Y}_1$,由定理 1 得到 $\mathbf{Y}_1=\mathbf{N}(\mathbf{N}(U_1^*))$, 而这仅当 $\mathbf{N}(U_1^*)=\{\mathbf{0}\}$ 时才可能,因而算子 U_1^* 实现了从 B-空间 $\mathbf{Y}_1^*$ 到空间 $U_1^*(\mathbf{Y}_1^*)$ 上的一对一的映射,而由(5)可知 $U_1^*(\mathbf{Y}_1^*)$ 也是 B-空间. 由定理 1. 2 的推论可知存在连续左逆算子 U_1^{*-1},而由定理 2 知有

$$U(\mathbf{X})=U_1(\mathbf{X})=\mathbf{Y}_1,$$

然后,再应用定理 1 就得到所需要的结果.

定理 3.* 如果 $U(\mathbf{X})$ 是闭集合,则

$$U^*(\mathbf{Y}^*)=\mathbf{N}^*(\mathbf{N}(U)), \tag{6}$$

即方程(2)有解的充分必要条件是泛函 f 在算子 U 的零集合上取值为零.

证. 先考察 $U(\mathbf{X})=\mathbf{Y}$ 的情形. 记 $\mathbf{X}_0=\mathbf{N}(U)=U^{-1}(\mathbf{0})$ 并引进商空间 $\overline{\mathbf{X}}=\mathbf{X}/\mathbf{X}_0$. 如同 1. 3 中那样,构造映空间 $\overline{\mathbf{X}}$ 到 $\mathbf{Y}$ 的算子 $\bar{U}$. 并且 $U=\bar{U}\varphi$, 其中 φ 是空间 $\mathbf{X}$ 到 $\overline{\mathbf{X}}$ 上的自然同态映照. 因为 $\overline{\mathbf{X}}$ 是

---

*) 算子 U 和 U_1 的不同在于象 $U(\mathbf{X})$ 和 $U_1(\mathbf{X})$ 位于不同的空间中, 因此共轭算子 U^* 和 U_1^* 作用在不同的空间中.

完备空间，算子$\bar{U}$是一对一的，所以，根据定理 1.2 的推论可知存在连续算子$\bar{U}^{-1}$. 因而，由定理 2 知$\bar{U}^*(\mathbf{Y}^*)=\overline{\mathbf{X}}^*$. 但$U^*=\varphi^*\bar{U}^*$，因而

$$U^*(\mathbf{Y}^*)=\varphi^*(\bar{U}^*(\mathbf{Y}^*))=\varphi^*(\overline{\mathbf{X}}^*).$$

于是就只须来证明等式

$$\varphi^*(\overline{\mathbf{X}}^*)=\mathbf{N}^*(\mathbf{X}_0).$$

取任意泛函$f\in\varphi^*(\overline{\mathbf{X}}^*)$. 设$\bar{f}\in\overline{\mathbf{X}}^*$使得$f=\varphi^*(\bar{f})$. 因为

$$f(x)=\bar{f}(\varphi(x)) \qquad (x\in\mathbf{X}),$$

而当$x\in\mathbf{X}_0$时$\varphi(x)=0$，所以$f\in\mathbf{N}^*(\mathbf{X}_0)$.

现将f看作$\mathbf{N}^*(\mathbf{X}_0)$中任意一个泛函，经过与在 1.3 中构造算子$\bar{U}$那样几乎一样的论证，我们即可推得泛函

$$\bar{f}(\bar{x})=\bar{f}(x) \qquad (x\in\bar{x}=\varphi(x))$$

有确切的定义并且是$\overline{\mathbf{X}}$中的连续线性泛函. 但这时

$$f=\varphi^*(\bar{f})\in\varphi^*(\overline{\mathbf{X}}^*).$$

在一般情形有可能$U(\mathbf{X})=\mathbf{Y}_1\neq\mathbf{Y}$，这时我们象在定理 3 的证明中一样引进从$\mathbf{X}$到$\mathbf{Y}_1$上的算子$U_1$. 在我们所考察的这种特殊情形下，算子$U_1$满足如下关系(注意到关系式(5))

$$U^*(\mathbf{Y}^*)=U_1^*(\mathbf{Y}^*)=\mathbf{N}^*(\mathbf{N}(U_1))=\mathbf{N}^*(\mathbf{N}(U)).$$

定理全部得证.

推论. 如果$U^*(\mathbf{Y}^*)$是闭集，则它也是$(*)$-弱闭集.

事实上，由定理 3 可推知$U(\mathbf{X})$是闭集，再根据定理 3*即可推知$U^*(\mathbf{Y}^*)$是$(*)$-弱闭集合.

我们再指出推论的一个特殊情形，即当U^*具有连续左逆的情形，这时，根据定理 1.1 的推论可知集合$U^*(\mathbf{Y}^*)$是闭集合.

2.5. 现在我们指出上述定理的几个推论.

定理 4. 方程(1)对任何$y\in\mathbf{Y}$存在唯一解的充分必要条件是方程(2)对任何$f\in\mathbf{X}^*$有唯一解.

定理 5. 设对任何 $y\in\mathbf{Y}$ 和 $f\in\mathbf{X}^*$ 方程(1)和(2)均可解，则一定存在连续(双边)逆算子 U^{-1} 和 U^{*-1}. 从而，方程(1)和(2)的解都是唯一的.

定理 6. 设当 $y=\mathbf{0}$ 和 $f=\mathbf{0}$ 时方程(1)和(2)仅有零解，这时，如果集合 $U(\mathbf{X})$ 和 $U^*(\mathbf{Y}^*)$ 有一个是闭集，则方程(1)和(2)对任意的 $y\in\mathbf{Y}$ 和 $f\in\mathbf{X}^*$ 有解，且解是唯一的.

注. 在定理 6 的条件下，集合 $U(\mathbf{X})$ 或 $U^*(\mathbf{Y}^*)$ 的闭性等价于存在连续左逆算子 U^{-1} 或 U^{*-1}.

2.6. 如果 $\mathbf{X}=\mathbf{H}$ 是 Hilbert 空间而 $\mathbf{Y}=\mathbf{X}$, 则由定理 6 可引导出这样的结果：自共轭算子 U 具有连续双边逆算子的充分必要条件是存在连续左逆算子 U^{-1}, 即对任意的 $x\in\mathbf{H}$, 有

$$\|U(x)\|\geqslant m\|x\| \qquad (m>0),$$

这个结果我们在第九章中从另外的考虑出发曾经得到过(参见定理 IX. 5. 3), 不过上述的结果是更为一般的形式.

由定理 5 还得到下述定理.

定理 7. 如果自共轭算子 U 将空间 $\mathbf{H}$ 映射到自身上，则存在连续逆算子 U^{-1}.

2.7. 我们以下述线性代数方程组

$$\begin{aligned} a_{11}\xi_1+a_{12}\xi_2+\cdots+a_{1m}\xi_m&=\eta_1,\\ a_{21}\xi_1+a_{22}\xi_2+\cdots+a_{2m}\xi_m&=\eta_2,\\ \cdots\cdots\cdots\cdots&\cdots\cdots\\ a_{n1}\xi_1+a_{n2}\xi_2+\cdots+a_{nm}\xi_m&=\eta_n \end{aligned} \tag{7}$$

为例来说明这一节中证明的定理. 上述方程组可以简写成一个方程

$$U(x)=y,$$

其中U是由矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

所确定的从m维赋范空间 $\mathbf{X}_m$ 到 n 维赋范空间 $\mathbf{Y}_n$*) 内的算子. 从 $\mathbf{Y}_n^*$到 $\mathbf{X}_m^*$内的共轭算子 U^* 由矩阵

$$A^*=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

来确定, 使得

$$f=U^*(g)$$

$$(f=(\varphi_1, \varphi_2, \cdots, \varphi_m)\in\mathbf{X}_m^*, \quad g=(\psi_1, \psi_2, \cdots, \psi_n)\in\mathbf{Y}_n^*)$$

表示

$$\varphi_k=\sum_{j=1}^{n}a_{jk}\psi_j \qquad (k=1, 2, \cdots, m).$$

因为 $U(\mathbf{X}_m)$是有限维线性集合, 故 $U(\mathbf{X}_m)$ 显然是闭的 (见 IV1.6), 因而由定理 1 得

$$U(\mathbf{X}_m)=\mathbf{N}(\mathbf{N}(U^*)).$$

但 $\mathbf{N}(U^*)$是由使得 $U^*(g)=\mathbf{0}$, 即使得

---

*) 在空间 $\mathbf{X}_m$ 和 $\mathbf{Y}_n$ 中用什么方式引进范数是没有实质性差别的, 为确定起见, 例如可认为 $X_m=\mathbf{l}_m^2$, $\mathbf{Y}_n=\mathbf{l}_n^2$ 且为实空间.

$$\sum_{j=1}^{n} a_{jk}\psi_j = 0 \qquad (k=1, 2, \cdots, m) \tag{8}$$

的泛函 $g=(\psi_1, \psi_2, \cdots, \psi_n)\in \mathbf{Y}_n^*$ 所组成的．因此 $y=(\eta_1, \eta_2, \cdots, \eta_n)\in U(\mathbf{X}_m)$ 等价于从(8)推出

$$g(y)=\sum_{j=1}^{n}\psi_j\eta_j=0.$$

换言之，方程组(7)有解的充分必要条件是(7)右端项构成的向量与以共轭矩阵为系数矩阵的齐次方程组的任一解均正交．

如果 $m=n$ 且(7)对任何右端项可解，则算子 U^* 有左逆(定理 2*)．不难看出，这时 $U^*(\mathbf{Y}_n^*)=\mathbf{X}_n^*$，因若不然则算子 U^* 将把 n 维空间映射到较低维的空间，这与存在左逆 U^{*-1} 矛盾*)．应用定理 4 即可知方程组(7)的解是唯一的．

利用定理 5 并经类似的论证可得逆命题：如果方程组(7)具有不多于一个解，则对任何右端项解均存在．

对于具有对称系数矩阵的方程组而言，这两个结果可直接从定理 7 推出来．

---

*）事实上，设集合 $U^*(\mathbf{Y}_n^*)$ 有维数 $n_0<n$，即存在元素 $f_1, f_2, \cdots, f_{n_0}\in U^*(\mathbf{Y}_n^*)$ 使得对于任意的 $f\in U^*(\mathbf{Y}_n^*)$

$$f=\sum_{k=1}^{n_0}c_kf_k.$$

对任意的 $g\in\mathbf{Y}_n^*$ 这时有 $g=\sum_{k=1}^{n_0}c_kg_k$，其中 $g_k=U^{*-1}(f_k)$，即 $\mathbf{Y}_n^*$ 的维数小于 n.

第十三章　第二类泛函方程

在本章中我们将研究形如

$$x-\lambda U(x)=y \tag{*}$$

的方程，其中U是从B-空间$\mathbf{X}$到自身的连续线性算子．我们把这样的方程叫作第二类方程，而算子U称作方程的核．这个术语来源于积分方程论．在积分方程中，我们将

$$x(s)-\lambda\int_a^b K(s,t)x(t)dt=y(s) \qquad (s\in[a,b])$$

叫第二类方程以区别于如下的第一类积分方程

$$\int_a^b K(s,t)x(t)dt=y(s) \qquad (s\in[a,b]).$$

尽管泛函方程(*)也可以形式地改写为“第一类”方程

$$T(x)=y \qquad (T=I-\lambda U)$$

的形式，但由于算子U较之算子T具有更好的性质，所以分出恒等算子能使我们比较充分地研究方程(*)．

§1. 具有紧核的方程

在本节中我们考察方程

$$x-U(x)=y \qquad (x,y\in\mathbf{X}) \tag{1}$$

及其共轭方程

$$g-U^*(g)=f \qquad (f,g\in\mathbf{X}^*), \tag{2}$$

并假设U是B-空间$\mathbf{X}$中的紧算子(从而U^*也是紧算子——定理IX.3.3)．引进记号$T=I-U$，其中I表示空间$\mathbf{X}$中的恒等算子．这时，方程(1)可以简写为

$$T(x)=y, \tag{1'}$$

又因为 $T^*=I^*-U^*$(见 IX. 3. 1)，I^* 是 $\mathbf{X}^*$ 中的恒等算子，故知方程(2)可写成

$$T^*(g)=f. \tag{2'}$$

1. 1. 首先来证明三个引理.

引理 1. $T(\mathbf{X})$是闭集合.

证. 记 $\mathbf{X}_0=\mathbf{N}(T)=T^{-1}(\mathbf{0})$并考察商空间 $\overline{\mathbf{X}}=\mathbf{X}/\mathbf{X}_0$ 和从空间 $\overline{\mathbf{X}}$ 到 $\mathbf{X}$ 的算子 $\overline{T}$(见 XII. 1. 3). 如同在 XII. 1. 3 中那样，用 φ 表示空间 $\mathbf{X}$ 到 $\overline{\mathbf{X}}$ 上的自然同态. 设 $\{y_n\}\subset T(\mathbf{X})$ 是收敛于 $y_0\in\mathbf{X}$ 的序列. 因为 $\overline{T}(\overline{\mathbf{X}})=T(\mathbf{X})$, 所以存在元素 $\bar{x}_n\in\overline{\mathbf{X}}$ 使得

$$y_n=\overline{T}(\bar{x}_n) \qquad (n=1, 2, \cdots).$$

其次, 我们再求得与 IV. 1. 8 中关系式(1)相应的元素 $x_n\in\mathbf{X}$, 即求得满足如下关系的 $x_n\in\mathbf{X}$:

$$\bar{x}_n=\varphi(x_n), \|\bar{x}_n\|\geqslant\frac{1}{2}\|x_n\| \quad (n=1, 2, \cdots). \tag{3}$$

我们来证明序列$\{\bar{x}_n\}$是有界的. 如若不然，则可找到子序列使得 $c_n=\|\bar{x}_n\|\to\infty$. 由(3)可知序列$\left\{\frac{x_n}{c_n}\right\}$有界, 所以可再找到子序列, 可以认为$\left\{U\left(\frac{x_n}{c_n}\right)\right\}$收敛. 例如可设$U\left(\frac{x_n}{c_n}\right)\to z$. 由于

$$T(x_n)=\overline{T}(\bar{x}_n)=y_n,$$

故可得

$$\frac{x_n}{c_n}=U\left(\frac{x_n}{c_n}\right)+T\left(\frac{x_n}{c_n}\right)=U\left(\frac{x_n}{c_n}\right)+\frac{y_n}{c_n}\to z,$$

因而

$$T(z)=\lim_{n\to\infty}T\left(\frac{x_n}{c_n}\right)=\lim_{n\to\infty}\left(\frac{y_n}{c_n}\right)=\mathbf{0},$$

也就是说 $z\in\mathbf{X}_0$. 但这时

$$\frac{\bar{x}_n}{c_n}=\varphi\left(\frac{x_n}{c_n}\right)\to\varphi(z)=\mathbf{0},$$

这是不可能的，因为对任何 $n=1,2,\cdots$ 均有 $\left\|\frac{\bar{x}_n}{c_n}\right\|=1$.

于是序列 $\{\bar{x}_n\}$ 有界，从(3)知序列 $\{x_n\}$ 也有界. 因而可以认为 $\{U(x_n)\}$ 是收敛序列. 例如设 $U(x_n)\to x$，则

$$x_n=T(x_n)+U(x_n)=y_n+U(x_n)\to y_0+x=x_0.$$

由此可得到

$$y_0=\lim_{n\to\infty}y_n=\lim_{n\to\infty}T(x_n)=T(x_0)\in T(\mathbf{X}),$$

这就是所要证明的.

引理 2. 集合序列

$$\mathbf{N}(T),\mathbf{N}(T^2),\mathbf{N}(T^3),\cdots,\mathbf{N}(T^n),\cdots$$

是递增的且只包含有限个不同的集合.

证. 结论的第一部分几乎是显然的；因为若 $x\in\mathbf{N}(T^n)$，则 $T^n(x)=\mathbf{0}$，故 $T^{n+1}(x)=\mathbf{0}$，即 $x\in\mathbf{N}(T^{n+1})$，$\mathbf{N}(T^n)\subset\mathbf{N}(T^{n+1})$.

为了证明第二部分，命 $\mathbf{X}_n=\mathbf{N}(T^n)$ 并来证明：若对某个 $n=1,2,\cdots$ 有 $\mathbf{X}_n=\mathbf{X}_{n+1}$，则亦有 $\mathbf{X}_{n+1}=\mathbf{X}_{n+2}$. 取 $x\in\mathbf{X}_{n+2}$ 这就是说

$$T^{n+2}(x)=T^{n+1}(T(x))=\mathbf{0},$$

因而 $T(x)\in\mathbf{X}_{n+1}=\mathbf{X}_n$. 这样就有

$$T^{n+1}(x)=T^n(T(x))=\mathbf{0},$$

即 $x\in\mathbf{X}_{n+1}$. 由此可见，$\mathbf{X}_{n+2}\subset\mathbf{X}_{n+1}$. 相反的包含关系总是成立的，因而 $\mathbf{X}_n=\mathbf{X}_{n+1}=\mathbf{X}_{n+2}=\cdots$.

现在假设对于每个 $n=1,2,\cdots$ 均有

$$\mathbf{X}_n\neq\mathbf{X}_{n+1}.$$

每个 $\mathbf{X}_n$ 是 $\mathbf{X}_{n+1}$ 的子空间，因此，由 IV.1.7 中的几乎垂直引理可知存在规格化元素 x_{n+1}，使得

$$\rho(x_{n+1},\mathbf{X}_n)>\frac{1}{2}\qquad(n=0,1,\cdots).\tag{4}$$

设 $m>n$. 考察元素

$$U(x_m)-U(x_n)=x_m-T(x_m)-[x_n-T(x_n)]=x_m-\tilde{x},$$

其中 $\tilde{x}=T(x_m)+x_n-T(x_n)$. 现证 $\tilde{x}\in \mathbf{X}_{m-1}$. 事实上,

$$T^{m-1}(\tilde{x})=T^m(x_m)+T^{m-1}(x_n)-T^m(x_n)=\mathbf{0},$$

这是因为 $x_n\in \mathbf{X}_n\subset \mathbf{X}_{m-1}$ 而 $x_m\in \mathbf{X}_m$ 之故.

注意到(4)式,即得

$$\|U(x_m)-U(x_n)\|=\|x_m-\tilde{x}\|>\frac{1}{2}\quad (m>n;\quad m,n=1,2,\cdots).\ (5)$$

但$\{x_n\}$是有界序列,由算子U的紧性, 从$\{U(x_n)\}$中可抽出收敛的子序列. 这个结果与(5)矛盾.

引理 3. 集合

$$T(\mathbf{X}),T^2(\mathbf{X}),\cdots,T^n(\mathbf{X}),\cdots \tag{6}$$

中仅有有限个是不相同的.

证. 和上一引理的证法大致相同, 所以不再进行很详细的推导了.

由引理 1,集合(6)都是闭的而且(6)构成递降序列. 显然, 从对某个 n 成立等式 $T^n(\mathbf{X})=T^{n+1}(\mathbf{X})$ 可推出

$$T^n(\mathbf{X})=T^{n+1}(\mathbf{X})=T^{n+2}(\mathbf{X})=\cdots,$$

因此,对这种情形引理就得到了证明.

假定 $T^n(\mathbf{X})\neq T^{n+1}(\mathbf{X})\ (n=0,1,\cdots)$,利用关于几乎垂直的引理(见IV.1.7),可构造序列$\{x_n\}$使得

$$\|x_n\|=1,\ x_n\in T^n(\mathbf{X}),\ \rho(x_n,T^{n+1}(\mathbf{X}))>\frac{1}{2}\quad (n=1,2,\cdots).\ (7)$$

设 $m>n$. 如同引理 2 那样,有如下等式

$$U(x_n)-U(x_m)=x_n-T(x_n)-[x_m-T(x_m)]=x_n-\tilde{x}.$$

但是

$$T(x_n)\in T^{n+1}(\mathbf{X}),\ x_m\in T^m(\mathbf{X})\subset T^{n+1}(\mathbf{X}),$$

$$T(x_m)\in T^{m+1}(\mathbf{X})\subset T^{n+1}(\mathbf{X}),$$

因此，$\tilde{x}=T(x_n)+x_m-T(x_m)\in T^{m+1}(\mathbf{X})$. 故由(7)推出

$$\|U(x_n)-U(x_m)\|=\|x_n-\tilde{x}\|>\frac{1}{2}\quad (m>n;\ m,n=1,2,\cdots),$$

这与算子U的紧性矛盾.

1.2. 用r表示使得等式$T^n(\mathbf{X})=T^{n+1}(\mathbf{X})$成立的$n$中最小的非负整数. 特别，如果$T(\mathbf{X})=\mathbf{X}=T^0(\mathbf{X})$，则命$r=0$.

再记$\mathbf{X}'=T^r(\mathbf{X})$, $\mathbf{X}''=\mathbf{N}(T^r)$.

下述定理表明了算子T(因而也是方程(1))的特征.

定理 1. a) 算子T将子空间$\mathbf{X}'$一对一的映射到自身.

b) 子空间$\mathbf{X}''$是有限维的，算子T将$\mathbf{X}''$映射到自身.

c) 每个元素$x\in\mathbf{X}$可唯一地表示成如下形式

$$x=x'+x''\qquad (x'\in\mathbf{X}',\ x''\in\mathbf{X}''),\tag{8}$$

并且存在常数$M>0$，使得

$$\|x'\|\leqslant M\|x\|,\ \|x''\|\leqslant M\|x\|.\tag{9}$$

d) 算子U可表示成

$$U=U'+U'',\tag{10}$$

其中U'和U''分别是从空间$\mathbf{X}$到$\mathbf{X}'$和到$\mathbf{X}''$的紧算子. 此外，算子$T'=I-U'$有双边连续逆算子且成立下式

$$U'U''=U''U'=\mathbf{0}.\tag{11}$$

证. a) 因为$\mathbf{X}'=T^r(\mathbf{X})$，故

$$T(\mathbf{X}')=T^{r+1}(\mathbf{X})=T^r(\mathbf{X})=\mathbf{X}'.$$

如果$T(x)=\mathbf{0}$，其中$x\in\mathbf{X}'$，取$n\geqslant r$使得满足引理2中的等式$\mathbf{N}(T^n)=\mathbf{N}(T^{n+1})$，故$x\in T^n(\mathbf{X})$，因而存在$\tilde{x}\in\mathbf{X}$使得$x=T^n(\tilde{x})$. 但这时$\mathbf{0}=T(x)=T^{n+1}(\tilde{x})$，故$\tilde{x}\in\mathbf{N}(T^{n+1})=\mathbf{N}(T^n)$，即

$$x=T^n(\tilde{x})=\mathbf{0}.$$

b) 我们有

$$T^r=(I-U)^r=I-U_1,$$

其中算子U_1是算子U的正次幂的线性组合．因而由定理 IX.2.2 算子U_1是紧的．因为对$x\in\mathbf{X}''$有$U_1(x)=x$，所以$\mathbf{X}''$中每个有界集均紧．根据定理 IV.1.3 可知$\mathbf{X}''$是有限维的．

当$r>0$时，集合$T(\mathbf{X}'')$显然是$\mathrm{N}(T^{r-1})\subset\mathrm{N}(T^r)=\mathbf{X}''$．而若$r=0$，则$\mathbf{X}''=\{\mathbf{0}\}$并且显然有包含关系$T(\mathbf{X}'')\subset\mathbf{X}''$．

c) 用T_0表示算子T仅在集$\mathbf{X}'$上考察时得到的算子．应用引理1于算子$T^r=I-U_1$可知$\mathbf{X}'$是闭集，因之是B-空间．因此，由定理 XII.1.2 可知映$\mathbf{X}'$到自身的一对一的算子T_0具有连续逆算子T_0^{-1}．

设x是$\mathbf{X}$中任一元素，命

$$x'=T_0^{-r}T^r(x),\ x''=x-x'=x-T_0^{-r}T^r(x). \tag{12}$$

显然，$x'\in\mathbf{X}'$，又因为

$$T^r(x'')=T^r(x)-T^rT_0^{-r}T^r(x)=T^r(x)-T^r(x)=\mathbf{0},$$

所以$x''\in\mathbf{X}''$，此即证明了将x表示成(8)式的可能性．

现在来证(8)式的表示还是唯一的．设$x=x_1'+x_1''$是元素x形如(8)式的另一种表示，于是$x_1'\in\mathbf{X}'$，$x_1''\in\mathbf{X}''$，故有

$$T^r(x)=T^r(x_1')+T^r(x_1'')=T^r(x_1').$$

但因为$x_1'\in\mathbf{X}'$，故$T^r(x_1')=T_0^r(x_1')$，所以

$$x_1'=T_0^{-r}T^r(x)=x',$$

这就证明了(8)式表示的唯一性．

根据(12)，从算子T_0^{-1}的连续性即推得式(9)．

d) 因为$U=I-T$，故对于$x\in\mathbf{X}'$有

$$U(x)=x-T(x)\in\mathbf{X}',$$

即算子U将$\mathbf{X}'$映射到自身．类似地可得$U(\mathbf{X}'')\subset\mathbf{X}''$．

对任意的$x\in\mathbf{X}$，命

$$U'(x)=U(x'),\quad U''(x)=U(x''), \tag{13}$$

其中 $x'\in\mathbf{X}'$, $x''\in\mathbf{X}''$ 是元素 x 的形如(8) 式中的对应分量. 由式(9)不难知 U' 和 U'' 是连续线性算子. 此外, 显然有 $U=U'+U''$ 而且 $U'(\mathbf{X})\subset\mathbf{X}'$, $U''(\mathbf{X})\subset\mathbf{X}''$. 另外, 显然有

$$U'(\mathbf{X}'')=U''(\mathbf{X}')=\{\mathbf{0}\}. \tag{14}$$

由这两个关系式可以得到 $U'U''=U''U'=\mathbf{0}$, 即(11)式成立.

算子 U'' 将空间 $\mathbf{X}$ 映射到有限维空间 $\mathbf{X}''$, $\mathbf{X}''$ 中的所有有界集均是紧的, 所以 U'' 是紧算子. 但 $U'=U-U''$, 由定理 IX. 2. 2 可知算子 U' 也是紧的.

最后, 我们来证明算子 $T'=I-U'$ 具有双边连续逆算子. 为此, 只须证明两点: 1° 由 $T'(x)=\mathbf{0}$ 可推知 $x=\mathbf{0}$, 2° $T'(\mathbf{X})=\mathbf{X}$. 设 $T'(x)=\mathbf{0}$. 将 x 表示成形如(8)的形式我们就得到

$$\mathbf{0}=T'(x)=x-U'(x)=x'-U(x')+x''=T(x')+x''.$$

因为 $T(x')\in\mathbf{X}'$, 故由零元素的形如(8)表示的唯一性可知

$$T(x')=x''=\mathbf{0},$$

而由 a) 可知 $x'=\mathbf{0}$. 因而 $x=x'+x''=\mathbf{0}$.

现在考察任一元素 $y\in\mathbf{X}$. 将 y 表示成(8)的形式:

$$y=y'+y''\,(y'\in\mathbf{X}',\,y''\in\mathbf{X}'')$$

并命

$$x=T_0^{-1}(y')+y''.$$

因为 $T_0^{-1}(y')\in\mathbf{X}'$, 则

$$U'(x)=U(T_0^{-1}(y'))$$

并且

$$\begin{aligned}T'(x)&=x-U'(x)=T_0^{-1}(y')-U(T_0^{-1}(y'))+y''\\&=TT_0^{-1}(y')+y''=y'+y''=y,\end{aligned}$$

于是 $T'(\mathbf{X})=\mathbf{X}$.

定理完全得证.

注. 设 m 表示使得等式 $\mathbf{N}(T^n)=\mathbf{N}(T^{n+1})$ 成立的 n 中最小

的非负整数，则这时 $m=r$.

事实上，取 $x\in\mathbf{N}(T^{r+1})$ 并将 x 表示成(8)，可得

$$\mathbf{0}=T^{r+1}(x)=T^{r+1}(x')+T^{r+1}(x'')=T^{r+1}(x'),$$

由 a) 知上式仅在 $x'=\mathbf{0}$ 时才可能．由此 $x=x''\in\mathbf{N}(T^r)$，所以

$$m\leqslant r.$$

其次，如果 $y=T^m(x)(x\in\mathbf{X})$，将 x 表示成(8)式的分解式，则得

$$y=T^m(x)=T^m(x')+T^m(x'')=T^m(x')=T^{m+1}(T_0^{-1}(x')),$$

于是 $y\in T^{m+1}(\mathbf{X})$，因而又应有 $r\leqslant m$.

这个注所指出的事实的特例包含在下述定理中.

定理 2. 方程(1)对任何 $y\in\mathbf{X}$ 均可解的充分必要条件是齐次方程

$$T(x)=\mathbf{0} \tag{15}$$

具有唯一的零解 $x=\mathbf{0}$.

事实上，方程(1)对任何 $y\in\mathbf{X}$ 的可解性即表示 $T(\mathbf{X})=\mathbf{X}$，亦即表示 $r=0$．方程(15)有唯一解等价于 $m=0$.

注. 用前一章 § 2 中的结果，不依靠定理 1，只利用 $T(\mathbf{X})$ 是闭集的事实也可证明此定理．建议读者独立地完成必要的论证.

1.3. 下述定理建立了方程(1)和方程(2)之间的联系.

定理 3. 集合 $\mathbf{N}(T)$ 和集合 $\mathbf{N}(T^*)$ 具有相同的有限维数.

证. 因为 $\mathbf{N}(T)\subset\mathbf{N}(T^r)=\mathbf{X}''$，而由定理 1 中的 b) 可知 $\mathbf{X}''$ 是有限维的，因此 $\mathbf{N}(T)$ 也是有限维的．因为 U^* 也是紧算子，类似地可知 $\mathbf{N}(T^*)$ 也是有限维的.

设 $\mathbf{N}(T)$ 的维数等于 n，$\mathbf{N}(T^*)$ 的维数等于 m．分别设 $x_1, x_2, \cdots, x_n$ 和 $g_1, g_2, \cdots, g_m$ 是 $\mathbf{N}(T)$ 和 $\mathbf{N}(T^*)$ 中的线性独立元素组.

因为元素 $x_1, x_2, \cdots, x_n$ 线性独立，对这一组元组应用双正交化定理(定理 V.7.4)，可知存在双正交泛函组 $f_1, f_2, \cdots, f_n$:

$$f_j(x_k)=\begin{cases}0, & j\neq k,\\ 1, & j=k\end{cases}\qquad (j,k=1,2,\cdots,n). \tag{16}$$

同样地,利用引理 III. 3. 1, 可求得 $y_1, y_2, \cdots, y_m$ 使得

$$g_j(y_k)=\begin{cases}0, & j\neq k,\\ 1, & j=k\end{cases}\qquad (j,k=1,2,\cdots,m). \tag{17}$$

先假定 $n<m$, 在空间$\mathbf{X}$中考察算子 $V=U+W$, 其中

$$W(x)=\sum_{k=1}^{n} f_k(x)y_k \qquad (x\in\mathbf{X}).$$

因为线性算子 W 将$\mathbf{X}$变到有限维空间, 所以 W 是紧算子. 由此可知V 也是紧算子. 考察方程

$$\tilde{T}(x)=x-V(x)=T(x)-\sum_{k=1}^{n} f_k(x)y_k=\mathbf{0}. \tag{18}$$

设 x_0 是(18)的某个解, 亦即

$$\tilde{T}(x_0)=T(x_0)-\sum_{k=1}^{n} f_k(x_0)y_k=\mathbf{0}. \tag{19}$$

由(19)式推得

$$g_s(T(x_0))-\sum_{k=1}^{n} f_k(x_0)g_s(y_k)=0 \qquad (s=1,2,\cdots,n). \tag{20}$$

注意到(17), 则(20)可表示成

$$g_s(T(x_0))-f_s(x_0)=0,$$

由于 $T^*(g_s)=\mathbf{0}$, 故

$$f_s(x_0)=0 \qquad (s=1,2,\cdots,n). \tag{21}$$

从(19)式得到 $T(x_0)=\mathbf{0}$, 即 $x_0\in\mathbf{N}(T)$, 这表明 x_0 可以表示成如下的形式

$$x_0=\sum_{k=1}^{n}\alpha_k x_k.$$

因从(16)知 $\alpha_s=f_s(x_0)$，则由(21)得到 $\alpha_s=0$，于是 $x_0=\mathbf{0}$. 于是方程(18)有唯一解. 按定理 2，相应的非齐次方程对任何右端项均可解. 特别是，方程

$$\tilde{T}(x)=T(x)-\sum_{k=1}^{n}f_k(x)y_k=y_{n+1}$$

有解. 设此解为 x^*. 另一方面，

$$g_{n+1}(T(x^*)-\sum_{k=1}^{n}f_k(x^*)y_k)$$

$$=T^*(g_{n+1})(x^*)-\sum_{k=1}^{n}f_k(x^*)g_{n+1}(y_k)=0$$

然而 $g_{n+1}(y_{n+1})=1$.

因此，应有 $m\leqslant n$.

用类似的论证可以说明 $m<n$ 是不可能的. 亦即，应在空间 $\mathbf{X}^*$ 中考察

$$T^*(g)-\sum_{k=1}^{m}g(y_k)f_k=\mathbf{0}$$

代替(18)式.

1.4. 综合上述各定理可得如下结果.

定理 4. 或者是方程(1)和方程(2)对任何右端项均可解且这时其解都是唯一的；或者是齐次方程

$$T(x)=0 \quad 和 \quad T^*(g)=0$$

分别具有数量相同的线性独立解 $x_1, x_2, \cdots, x_n$ 和 $g_1, g_2, \cdots, g_n$，其中 n 是一个有限的正整数. 在这种情况下，方程(1)(相应地，方程(2))有解的充分必要条件是

$$g_k(y)=0 \qquad (k=1,2,\cdots,n)$$

(相应地，$f(x_k)=0 \qquad (k=1,2,\cdots,n)$).

这时，方程(1)和(2)的通解分别具有如下形式

$$x = x^* + \sum_{k=1}^{n} c_k x_k,$$

$$g = g^* + \sum_{k=1}^{n} c_k g_k,$$

其中 x^* 和 g^* 分别是方程(1)和(2)的某个解，而 $c_1, c_2, \cdots, c_n$ 是任意常数.

对方程(1)和(2)应用定理 XII. 2. 3 和定理 XII. 2. 3*(由引理 1 可知这些定理的条件成立)即得定理的第二部分（但也可以直接证明）.

由于和大家所熟知的积分方程论中定理的类似性，所以这个定理称为 Fredholm 择一律.

§2. 关于复赋范空间

在下面(参见 §3)，我们将要在复空间中来考察方程

$$x - \lambda U(x) = y,$$

因此，也就很自然地认为其中的 λ 是复参数. 为此，我们在此要引进关于复空间的一些准备知识，这些概念以实空间中的相应概念为其特款.

2.1. 设 **Z** 是复赋范空间. 如果在 **Z** 上定义一个将 **Z** 变换到自身的满足如下条件的算子 C:

1. $C(\lambda_1 z_1 + \lambda_2 z_2) = \bar{\lambda}_1 C(z_1) + \bar{\lambda}_2 C(z_2) \quad (z_1, z_2 \in \mathbf{Z})$,
2. $C^2(z) = z \quad (z \in \mathbf{Z})$,
3. $\|C(z)\| = \|z\|$,

则称 **Z** 具有实核，称 C 为对合算子.

所有使得 $C(z) = z$ 的 z 的集合叫做空间 **Z** 的实核并记为 Re **Z**，此集合中的元素叫作实的.

设 $z\in\mathbf{Z}$，元素 $x=\frac{1}{2}[z+C(z)]$ 叫 z 的实部，记为 $x=\operatorname{Re}z$. 元素 $y=\frac{1}{2i}[z-C(z)]$ 叫 z 的虚部，记为 $y=\operatorname{Im}z$. 因为

$$C(x)=\frac{1}{2}[C(z)+C^2(z)]=x,$$

$$C(y)=-\frac{1}{2i}[C(z)-C^2(z)]=y,$$

故 $x=\operatorname{Re}z$ 和 $y=\operatorname{Im}z$ 都是实元素. 显然,

$$z=x+yi, \tag{1}$$

$$C(z)=x-yi.$$

由于后一等式，我们将 $C(z)$ 称作是 z 的共轭并引用通常的记号:

$$C(z)=\bar{z}.$$

如果将 z 表示成具有实分量 x 和 y 的形如(1)的形式，则必有

$$x=\operatorname{Re}z \text{ 而 } y=\operatorname{Im}z.$$

事实上,

$$\bar{z}=\bar{x}+\overline{yi}=x-yi,$$

所以

$$x=\frac{1}{2}(z+\bar{z})=\operatorname{Re}z, \qquad y=\frac{1}{2i}(z-\bar{z})=\operatorname{Im}z.$$

因此这就表明表示(1)是唯一的，元素 x 和 y 由元素 z 唯一确定.

空间 Z 的实核 X 是实的赋范空间，如果原来的空间 Z 是完备的，则 X 就是实的 B-空间.

事实上，X 中元素按实系数的线性组合仍是 X 中的元素. 由于 Z 是赋范空间，故可推知 X 也满足赋范空间的公理. 假定 Z 是完备的，今证 X 也是完备的. 设 $\{x_n\}$ 是 X 中元素的自收敛序列. 在空间 Z 中考察这个元素序列，由于 Z 完备，故存在极限

$$\lim_{n\to\infty}x_n=z\in\mathbf{Z}.$$

由条件 1 知有

$$\overline{x_n - z} = \overline{x_n} - \bar{z},$$

所以

$$\|x_n - z\| = \|\overline{x_n - z}\| = \|x_n - \bar{z}\|,$$

由此可知 $x_n \to \bar{z}$. 由极限的唯一性可知 $\bar{z} = z$, 亦即 $z \in \mathbf{X}$.

前面考察过的所有的具体实空间都是其相应的复空间的实核. 例如, 实的 $\mathrm{C}(K)$ 就是复空间 $\mathrm{C}_C(K)$ 的实核. 这里, 对合算子就是复共轭算子 $x(t) \to \overline{x(t)}$. 对空间 $\mathbf{L}^p$, $\mathbf{V}$, $\mathbf{c}_0$ 等等也可类似地讨论.

一般来说, 任何一个实空间 $\mathbf{X}$ 均可视为某个复空间的实核, 即可视作空间 $\mathbf{Z}$ 的实核, $\mathbf{Z}$ 中的元素乃是 $\mathbf{X}$ 中元素的有序偶

$$z = (x, y) \qquad (x, y \in \mathbf{X}),$$

且对这些元素定义下述运算规则:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2),$$
$$\lambda(x, y) = (\alpha x - \beta y, \beta x + \alpha y)(\lambda = \alpha + \beta i, \alpha, \beta \text{ 为实数}),$$
$$\|(x, y)\| = \max_\theta \|x\cos\theta + y\sin\theta\|^{*)}. \tag{2}$$

为了确信 $\mathbf{X}$ 是空间 $\mathbf{Z}$ 的实核, 只要令

$$C((x, y)) = (x, -y).$$

这时, 空间 $\mathbf{Z}$ 中的实元素将仅仅是所有形如 $(x, \mathbf{0})$ 的元素. 因为

$$\alpha(x_1, \mathbf{0}) + \beta(x_2, \mathbf{0}) = (\alpha x_1 + \beta x_2, \mathbf{0}),$$
$$\|(x, \mathbf{0})\| = \max_\theta \|x\cos\theta\| = \|x\|,$$

所以, 将元素 $(x, \mathbf{0})$ 和 x 等同看待即得到所要的结果. 按照这样的看法, 我们即可用普通的记号 $x + yi$ 来代替 (x, y). 空间 $\mathbf{Z}$ 叫作 $\mathbf{X}$ 的复化空间.

---

*) 当然, 具体空间中的范数可以用另外的办法引入. 例如, 若 $\mathbf{X} = \mathbf{L}^p$, 则复空间 $\mathbf{L}_C^p$ 中通常的范数与此范数等价但并不相等.

2.2. 设$\mathbf{Z}$和$\mathbf{W}$是分别具有实核$\mathbf{X}=\mathrm{Re}\,\mathbf{Z}$和$\mathbf{Y}=\mathrm{Re}\,\mathbf{W}$的复空间. 从$\mathbf{Z}$到$\mathbf{W}$的连续线性算子$\tilde{U}$称作实算子，乃是指$\tilde{U}$将$\mathbf{Z}$中的实元素变成$\mathbf{W}$中的实元素，即指$\tilde{U}(\mathbf{X})\subset\mathbf{Y}$. 因此，实算子就诱导了一个从实空间$\mathbf{X}$到实空间$\mathbf{Y}$的连续线性算子。

反之，如果U是从$\mathbf{X}$到$\mathbf{Y}$的连续线性算子而$\mathbf{X}=\mathrm{Re}\mathbf{Z}$和$\mathbf{Y}=\mathrm{Re}\mathbf{W}$，命

$$\tilde{U}(z)=\tilde{U}(x+yi)=U(x)+iU(y)\qquad(z=x+yi),$$

则得到从复空间$\mathbf{Z}$到复空间$\mathbf{W}$的连续线性算子. 显然，在$\mathbf{X}$上算子$\tilde{U}$与U一样. 算子$\tilde{U}$叫算子U的复扩张.

我们指出，成立下面的不等式

$$\|U\|\leqslant\|\tilde{U}\|\leqslant 2\|U\|. \tag{3}$$

不等式的前一半是显然的. 注意到$\|\mathrm{Re}z\|\leqslant\|z\|$*)，则(3)的后一半可由下列一串不等式推出

$$\|\tilde{U}(z)\|=\|U(x)+iU(y)\|\leqslant\|U(x)\|+\|U(y)\|$$
$$\leqslant\|U\|(\|x\|+\|y\|)\leqslant 2\|U\|\,\|z\|.$$

在许多情形下可以证明等式$\|\tilde{U}\|=\|U\|$成立. 例如，若在$\mathbf{Z}$和$\mathbf{W}$中用(2)引入范数，则

$$\|\tilde{U}(z)\|=\max_{\theta}\|U(x)\cos\theta+U(y)\sin\theta\|$$
$$\leqslant\|U\|\max_{\theta}\|x\cos\theta+y\sin\theta\|=\|U\|\,\|z\|,$$

因而$\|\tilde{U}\|\leqslant\|U\|$，再由(3)式，即得$\|\tilde{U}\|=\|U\|$.

其次，如果对任意的$\varepsilon>0$，存在规格化实元素x使得

$$\|\tilde{U}(x)\|>\|\tilde{U}\|-\varepsilon,$$

则等式$\|\tilde{U}\|=\|U\|$也成立，这是因为此时有

$$\|\tilde{U}\|\leqslant\|\tilde{U}(x)\|+\varepsilon=\|U(x)\|+\varepsilon\leqslant\|U\|+\varepsilon.$$

---

*) $\|\mathrm{Re}z\|=\frac{1}{2}\|z+\bar{z}\|\leqslant\frac{1}{2}[\|z\|+\|\bar{z}\|]=\|z\|$.

例如，当 $\mathbf{Z}$ 和 $\mathbf{W}$ 是上面所指的几个函数空间而 $\tilde{U}$ 是具有实核 $K(s,t)$ 的积分算子

$$w=\tilde{U}(z),\ w(s)=\int_a^b K(s,t)x(t)\,dt$$

时，则此条件成立.

当 $\mathbf{Z}$ 和 $\mathbf{W}$ 是序列空间而 $\tilde{U}$ 是由实矩阵所确定的算子时则此条件亦成立.

2.3. 如果实算子 $\tilde{U}$ 是紧算子，则由 $\tilde{U}$ 所诱导的从空间 $\mathbf{X}=\mathrm{Re}\mathbf{Z}$ 到空间 $\mathbf{Y}=\mathrm{Re}\mathbf{W}$ 的算子 U 也是紧算子. 反之亦真，即当 U 是从 $\mathbf{X}$ 到 $\mathbf{Y}$ 的紧算子时，则其复扩张 $\tilde{U}$ 也是紧算子.

为了证明这个事实，只须引用如下结论：序列 $\{z_n\}$ 收敛等价于序列 $\{\mathrm{Re}z_n\}$ 和 $\{\mathrm{Im}z_n\}$ 均收敛.

2.4. 如果 $\mathbf{Z}$ 是具有实核的复空间，则其共轭空间 $\mathbf{Z}^*$ 也具有实核.

事实上，设 $f\in\mathbf{Z}^*$，定义对合算子如下：命 $C(f)=\bar{f}$:

$$\bar{f}(z)=\overline{f(\bar{z})}\qquad (z\in\mathbf{Z}). \tag{4}$$

首先证明 $\bar{f}$ 是连续线性泛函. 事实上，

$$\begin{aligned}\bar{f}(\lambda z_1+\mu z_2)&=\overline{f(\overline{\lambda z_1+\mu z_2})}=\overline{f(\bar{\lambda}\bar{z}_1+\bar{\mu}\bar{z}_2)}\\&=\overline{\bar{\lambda}f(\bar{z}_1)+\bar{\mu}f(\bar{z}_2)}=\lambda\bar{f}(z_1)+\mu\bar{f}(z_2)\end{aligned}$$

以及

$$|\bar{f}(z)|=|f(\bar{z})|\leqslant\|f\|\|\bar{z}\|=\|f\|\|z\|. \tag{5}$$

现在来验证对合定义中的条件 1—3 均成立. 利用泛函乘复数的规则，则得

$$\begin{aligned}\overline{(\lambda f_1+\mu f_2)}(z)&=\overline{(\lambda f_1+\mu f_2)(\bar{z})}=\overline{\bar{\lambda}f_1(\bar{z})+\bar{\mu}f_2(\bar{z})}\\&=\lambda\overline{f_1(\bar{z})}+\mu\overline{f_2(\bar{z})}=\lambda\bar{f}_1(z)+\mu\bar{f}_2(z)\\&=(\bar{\lambda}\bar{f}_1+\bar{\mu}\bar{f}_2)(z),\\ \bar{\bar{f}}(z)&=\overline{\bar{f}(\bar{z})}=\overline{\overline{f(\bar{\bar{z}})}}=f(z)\end{aligned}$$

最后，从(5)推出不等式 $\|\bar{f}\|\leqslant\|f\|$. 另一方面，$\|f\|=\|\bar{\bar{f}}\|\leqslant\|\bar{f}\|$，因

此$\|\bar{f}\|=\|f\|$.

用$\mathbf{X}$表示空间$\mathbf{Z}$的实核，我们来证明：如果φ是实空间$\mathbf{X}$中的线性泛函，则φ的复扩张f属于穷举了形如f这样泛函的空间$\mathbf{Z}^*$的实核．这两个结论不难从共轭泛函的定义推出来．事实上，如果f是泛函$\varphi\in\mathbf{X}^*$的复扩张，则

$$\bar{f}(z)=\overline{f(\bar{z})}=\overline{\varphi(x)-\varphi(y)i}=\varphi(x)+\varphi(y)i=f(z)$$
$$(z=x+yi\in\mathbf{Z}).$$

反之，如果泛函f是空间$\mathbf{Z}^*$中的实元素，即如果有$\bar{f}=f$，则由(4)，这就表明$\overline{f(\bar{z})}=f(z)$．特别，如果$z=x\in\mathbf{X}$则

$$\overline{f(x)}=f(x),$$

亦即f是实的.

已证得的结果可以叙述成如下的形式：共轭空间的实核是已知空间实核的共轭空间.

2.5. 自然期望实算子的共轭算子仍是实算子．设$\tilde{U}$是从空间$\mathbf{Z}$到$\mathbf{W}$的实算子而U是由$\tilde{U}$所诱导出来的从空间$\mathbf{X}$到空间$\mathbf{Y}$的算子，其中$\mathbf{X}=\mathrm{Re}\mathbf{Z}$，$\mathbf{Y}=\mathrm{Re}\mathbf{W}$．今证$\tilde{U}^*$是实算子而由它所诱导出的从$\mathbf{Y}^*$到$\mathbf{X}^*$的算子就是算子$U^*$．后一个论断也就是说：如果$\varphi$和$\psi$分别是空间$\mathbf{X}$和$\mathbf{Y}$中的泛函，而且

$$\varphi=U^*(\psi), \tag{6}$$

则

$$f=\tilde{U}^*(g), \tag{7}$$

其中f和g分别是泛函φ和ψ的复扩张；反之，如果实泛函$f\in\mathbf{Z}^*$和$g\in\mathbf{W}^*$由(7)所联系，则由它们所诱导出来的泛函$\varphi\in\mathbf{X}^*$和$\psi\in\mathbf{Y}^*$由(6)所联系.

$\tilde{U}^*$是实算子的证明很简单：如果$g\in\mathbf{W}^*$是实泛函，亦即如果$\bar{g}=g$而$f=\tilde{U}^*(g)$，则

$$\bar{f}(z)=\overline{f(\bar{z})}=\overline{g(\tilde{U}(\bar{z}))}=\overline{g(\overline{\tilde{U}(z)})}=g(\tilde{U}(z))=f(z).$$

其次，设 $\varphi=U^*\psi$，则($x=\mathrm{Re}z$, $y=\mathrm{Im}z$)

$$f(z)=\varphi(x)+i\varphi(y)=\psi(U(x))+i\psi(U(y))$$
$$=g(U(x)+iU(y))=g(\tilde{U}(z)),$$

于是 $f=\tilde{U}^*(g)$. 从(7)推出(6)，证明同样是很简单的.

§3. 谱

3.1. 在这一节和下一节中，我们将研究依赖于复参数 λ(或 μ)的方程

$$T_\lambda(x)=x-\lambda U(x)=y, \tag{1}$$

或方程

$$T'_\mu(x)=\mu x-U(x)=y \tag{1'}$$

的性态. 这里和后面总假定 U 是复 B-空间 $\mathbf{X}$ 中的连续线性算子(如果 $\mathbf{X}$ 是实空间，则考察 U 的复扩张). 我们只所以要考察这两种方程，是因为方程(1)在积分方程论中有用(在§6中将给出对积分方程论的应用)而方程(1′)在抽象泛函分析中研究算子 U 的谱性质时有用.

对方程(1)的可解性而言，复平面被分成了两个集合，一个是对任何右端项 $y\in\mathbf{X}$ 使得方程(1)具有唯一解的 λ 的集合 $\pi(U)$(因而算子 T_λ 有连续逆(参见XII.1.3))；另一个是复平面上除去 $\pi(U)$ 之外的集合 $\chi(U)$. 集合 $\pi(U)$ 中的点叫做算子 U 的非奇异值，集合 $\chi(U)$ 叫作算子 U 的特征集合.

类似地，对方程(1′)引进集合 $\rho(U)$，它是使得方程(1′)对任何右端都有单值解的 μ 的集合. $\rho(U)$ 的余集记为 $\sigma(U)$. 集 $\rho(U)$ 中的点叫作算子 U 的正则值而 $\rho(U)$ 本身叫作算子 U 的豫解集合；集合 $\sigma(U)$ 叫作算子 U 的谱.

如果对某个 λ，齐次方程

$$T_\lambda(x)=x-\lambda U(x)=0 \tag{2}$$

具有非零解，则 λ 叫作算子 U 的特征值. 显然，特征值全体的集合 $\chi_0(U)$ 包含在集合 $\chi(U)$ 之中. 方程(2)的每一个解叫作对应于给定特征值的特征元或特征向量. 集合 $\bigcup_{n=1}^{\infty}\mathbf{N}(T_\lambda^n)$ 叫作根子空间，这个空间的维数(有限的或无限的)叫做特征值 λ 的重数. 各种不同集合 $\mathbf{N}(T_\lambda^n)(n=1,2,\cdots)$ 的个数 r 叫作特征值的秩(参见 XIII.1.1 引理 2).

如果代替方程(2)而来考察方程(1′)的齐次方程

$$T'_\mu(x)=\mu x-U(x)=\mathbf{0}, \tag{2'}$$

则我们就得到特征值以及对应于给定特征值的特征元(向量)和根子空间等概念，这些概念对 Hilbert 空间中的算子已经给出定义(IX.4.1).

注意，如果 U 是 Hilbert 空间中的自共轭算子，而 μ 是其特征值，则 μ 的秩 $r=1$，即

$$\mathbf{N}(T'_\mu)=\mathbf{N}(T'^2_\mu)=\cdots=\mathbf{N}(T'^n_\mu)=\cdots. \tag{3}$$

因此，在这种情况下根子空间就是 $\mathbf{N}(T'_\mu)$，即由算子 U 的所有特征向量组成.

现在来证明关系式(3). 由于自共轭算子的特征值是实的，因此算子 T'_μ 及其所有的幂次均是自共轭算子. 为简单计，略去下标 μ 并取 $n=2^k$，则对于 $x\in\mathbf{N}(T'^{2^k})$ 有

$$(T'^{2^{k-1}}x, T'^{2^{k-1}}x)=(T'^{2^k}x, x)=0,$$

由此

$$T'^{2^{k-1}}x=\mathbf{0}.$$

继续如上过程. 最后得到等式 $T'x=\mathbf{0}$，即 $x\in\mathbf{N}(T')$，从而 $\mathbf{N}(T')\supset\mathbf{N}(T'^{2^k})$ 相反的包含关系对任何算子都是成立的，因而 $\mathbf{N}(T')=\mathbf{N}(T'^{2^k})$ $(k=1,2,\cdots)$. 其次，若取任意的 $n=2,3,\cdots$ 并使 k 满足 $n\leqslant 2^k$ 的关系，则得到明显的包含关系 $\mathbf{N}(T')\subset\mathbf{N}(T'^n)\subset\mathbf{N}(T'^{2^k})$. 由此可见 $\mathbf{N}(T')=\mathbf{N}(T'^n)$.

我们来指出同一个算子U的谱和特征集合之间的简单联系. 易见,如果$\lambda\in\chi(U)$,则$\mu=\frac{1}{\lambda}\in\sigma(U)$,反之亦真. 算子$U$的特征值$\lambda$和$\mu$之间也有同样的关系.这时重要的是要注意到对应于特征值λ的特征元也是对应于特征值$\mu=\frac{1}{\lambda}$的特征元,反之亦然. 其次,因当$\lambda\mu=1$时有$\mathbf{N}(T_\lambda^n)=\mathbf{N}(T_\mu'^n)(n=1,2,\cdots)$,所以上面所指出的事实可以推广到特征子空间上去. 因此就不必要象前面那样在术语中把对应于特征值λ的特征元和对应于特征值μ的特征元从概念上加以区分了.

上面所指出的特征集合和谱之间的关系使我们有可能视其方便从这些平行的但实质上是等价的概念中选择一个来研究, 而只是在非常例外的情况下才同时讨论两者.

现在来指出从上面诸概念导出的某些简单结论.

I. 关系$\lambda\in\pi(U)$等价于存在连续双边逆算子

$$B_\lambda=T_\lambda^{-1}=(I-\lambda U)^{-1}.$$

参见 XII. 1. 3 定理 XII. 1. 2 的推论.

II. 非奇异值的集合是开集,因而特征集合是闭集.

这可由下述定理推出: 如果某个算子有连续的逆算子, 则依范数与它逼近的算子也有连续逆算子(V. 4. 6). 在我们这个情形

$$\|T_\lambda-T_{\lambda'}\|=|\lambda-\lambda'|\,\|U\|,$$

于是,如果T_λ^{-1}存在,则对于充分小的$\lambda-\lambda'$而言$T_{\lambda'}^{-1}$也存在.

III. 圆$|\lambda|<\frac{1}{\|U\|}$包含在集合$\pi(U)$之中; 因而, 谱$\sigma(U)$整个位于圆$|\mu|\leqslant\|U\|$之内.

为了证明这一点, 只要利用 Banach 逆算子定理(V. 4. 5) 即可.

IV. 集合$\chi(U)$和$\chi(U^*)$相对于实轴对称地分布.

事实上(IX.3.1)

$$T_\lambda^* = [I - \lambda U]^* = I^* - \bar{\lambda} U^*.$$

而由定理 XII.2.4, 算子 T_λ^{-1} 和 $(T_\lambda^*)^{-1}$ 同时存在*).

V. 如果 **X** 是具有实核的空间, 而 U 是实算子, 则集合 $\chi_0(U)$ 相对于实轴是对称的. 此外, 如果 $\lambda \in \chi_0(U)$ 而 z 是对应于 λ 的特征元, 则特征值 $\bar{\lambda}$ 将对应于特征元 $\bar{z}$.

事实上,

$$\bar{z} - \bar{\lambda} U(\bar{z}) = \overline{z - \lambda U(z)},$$

由此推知等式 $z - \lambda U(z) = \mathbf{0}$ 和 $\bar{z} - \bar{\lambda} U(\bar{z}) = \mathbf{0}$ 等价.

注. 在 IX.5.3 中已经给出了 Hilbert 空间中自共轭算子谱的定义. 在那里所证明的定理 IX.5.3 表明第九章的定义和这里的定义是等价的.

3.2. 当 U 是紧算子时, 特征集合的结构可以揭示得相当充分.

定理 1. 如果 U 是紧算子, 则

a) 特征集合仅仅包含特征值, 亦即 $\chi(U) = \chi_0(U)$; 这时每个特征值都具有有限的重数;

b) 对于任意的 $r > 0$, 圆 $|\lambda| \leqslant r$ 中只可能包含有限个特征值;

c) 如果 $\lambda_1 \in \chi(U)$, $\lambda_2 \in \chi(U^*)$ 而且 $\lambda_1 \neq \bar{\lambda}_2$, 又若 $x_1 \in \mathbf{X}$ 和 $g_2 \in \mathbf{X}^*$ 是分别对应于 λ_1 和 λ_2 的特征元, 则

$$g_2(x_1) = 0.$$

证. a) 由定理 XII.1.4, 如果 $\lambda \in \pi(U)$, 则齐次方程(2)具有非零解. 由引理 2(XIII.1.1)可知特征子空间是有限维的. 事实上, 由这个引理, 存在 n_0 使得 $\mathbf{N}(T_\lambda^n) = \mathbf{N}(T_\lambda^{n_0})$ $(n \geqslant n_0)$. 因此, 在这种情况下特征子空间是 $\mathbf{N}(T_\lambda^{n_0})$. 但

$$T_\lambda^{n_0} = (I - \lambda U)^{n_0} = I - \tilde{U},$$

---

*) T_λ^* 应理解为 $(T_\lambda)^*$.

其中 $\tilde{U}=\sum_{k=1}^{n_0}(-1)^{k-1}C_{n_0}^k\lambda^k U^k$ 显然是紧算子. 因此,由我们已经引用过的定理 XII. 1. 4 可知齐次方程

$$T_\lambda^{n_0}(x)=\mathbf{0}$$

的解的集合构成有限维子空间,而 $\mathbf{N}(T_\lambda)\subset\mathbf{N}(T_\lambda^{n_0})$.

b) 用反证法. 若不然, 即假设在圆 $|\lambda|\leqslant r$ 内包含特征值的无限集合. 从这个集合中找出不相同的特征值的序列 $\{\lambda_n\}$: $|\lambda_n|\leqslant r(n=1,2,\cdots)$. 设 $\{x_n\}$ 是其相应的特征元序列:

$$x_n-\lambda_n U(x_n)=\mathbf{0}\qquad(n=1,2,\cdots).\tag{4}$$

我们用归纳法来证明对任何 $n=1,2,\cdots$, 元素 $x_1,x_2,\cdots,x_n$ 均线性独立. 对 $n=1$, 这是显然的. 设对某个 $n\geqslant1$ 结论已经成立. 现证对于元素 $x_1,x_2,\cdots,x_n,x_{n+1}$ 结论亦成立. 若不然,将有

$$x_{n+1}=\sum_{k=1}^n\alpha_k x_k,$$

由此,根据(4)

$$\frac{x_{n+1}}{\lambda_{n+1}}=\sum_{k=1}^n\frac{\alpha_k}{\lambda_k}x_k.$$

上两式联立,得到

$$\sum_{k=1}^n\left(1-\frac{\lambda_{n+1}}{\lambda_k}\right)\alpha_k x_k=\mathbf{0}.$$

因为 $1-\frac{\lambda_{n+1}}{\lambda_k}\neq0(k=1,2,\cdots,n)$, 这样, 则元素 $x_1,x_2,\cdots,x_n$ 是线性相关的,此与归纳假定矛盾.

构造集合 $\mathbf{X}_n=\mathscr{L}(\{x_1,x_2,\cdots,x_n\})$. 因按已证明的, $\mathbf{X}_{n+1}\neq\mathbf{X}_n$ 所以按几乎垂直引理(参见 IV. 1. 7)可找到元素 $y_1,y_2,\cdots,y_n$ 使得

$$y_n\in\mathbf{X}_n,\quad\|y_n\|=1,\quad\rho(y_{n+1},\mathbf{X}_n)>\frac{1}{2}\qquad(n=1,2,\cdots).\tag{5}$$

如果 $x\in\mathbf{X}_n$, 即若

$$x=\sum_{k=1}^{n}\beta_k x_k,$$

则

$$U(x)=\sum_{k=1}^{n}\frac{\beta_k}{\lambda_k}x_k\in\mathbf{X}_n.$$

因此,

$$T_{\lambda_n}(x)=x-\lambda_n U(x)=\sum_{k=1}^{n}\beta_k\left(1-\frac{\lambda_n}{\lambda_k}\right)x_k\in\mathbf{X}_{n-1}$$

(因为 x_n 的系数等于零).

设 $m>n$, 考察下式

$$\begin{aligned}U(\lambda_m y_m)-U(\lambda_n y_n)&=y_m-[T_{\lambda_m}(y_m)+U(\lambda_n y_n)]\\&=y_m-\tilde{y}.\end{aligned}$$

由已经证明的可知 $T_{\lambda_m}(y_m)\in\mathbf{X}_{m-1}$ 及 $U(\lambda_n y_n)\in\mathbf{X}_n\subset\mathbf{X}_{m-1}$, 因而

$$\tilde{y}=T_{\lambda_m}(y_m)+U(\lambda_n y_n)\in\mathbf{X}_{m-1}.$$

由于(5),

$$\|U(\lambda_m y_m)-U(\lambda_n y_n)\|=\|y_m-\tilde{y}\|>\frac{1}{2}.$$

这和算子 U 的紧性矛盾, 因为序列 $\{\lambda_n y_n\}$ 有界 ($\|\lambda_n y_n\|\leqslant r$).

c) 由于 $\lambda_1 U(x_1)=x_1$, $\lambda_2 U^*(g_2)=g_2$, 因而

$$\begin{aligned}g_2(x_1)&=g_2(\lambda_1 U(x_1))=\lambda_1 U^*(g_2)(x_1)=\lambda_1\left(\frac{g_2}{\lambda_2}\right)(x_1)\\&=\frac{\lambda_1}{\lambda_2}g_2(x_1),\end{aligned}$$

由于 $\lambda_1\neq\bar{\lambda}_2$, 所以仅当 $g_2(x_1)=0$ 时上式才能成立.

最后指出, 如果 U 是无限维空间 $\mathbf{X}$ 中的紧算子, 则零点属于算子 U 的谱.

§4. 豫 解 式

4.1. 在这里我们继续研究方程

$$x-\lambda U(x)=y, \tag{1}$$

然而与上一节不同的是我们现在感兴趣的是它的单值可解性.

设 $\lambda\neq 0$ 是算子 U 的特征值. 由关系式

$$I+\lambda B_\lambda=(I-\lambda U)^{-1} \tag{2}$$

所确定的算子 B_λ 称作算子 U 的豫解式. 当 $\lambda=0$ 时令 $B_0=U$.

如果考察谱和正则值的集合,则代替 B_λ 而去考察算子

$$R_\mu=(\mu I-U)^{-1} \tag{3}$$

较为方便,算子 R_μ 对算子 U 的所有正则值都有意义. 我们同样把算子 R_μ 叫豫解式. 这两个概念不会混淆,因为从上下文总可以明白讲的是什么样的豫解式. 此外, 对不同的豫解式将用不同的记号来区分. 豫解式 B_λ 常在积分方程论中遇到(称为 Fredholm 豫解式),豫解式 R_μ 常在泛函分析中遇到(见 3.1).

如果 $\mu\neq 0$ 则显然

$$R_\mu=\frac{1}{\mu}I+\frac{1}{\mu^2}B_{\frac{1}{\mu}}. \tag{4}$$

反之,由等式

$$(I+\lambda B_\lambda)(I-\lambda U)=(I-\lambda U)(I+\lambda B_\lambda)=I$$

得到

$$B_\lambda=U(I-\lambda U)^{-1}=(I-\lambda U)^{-1}U. \tag{5}$$

因而,当 $\lambda\neq 0$ 时

$$B_\lambda=\frac{1}{\lambda}UR_{\frac{1}{\lambda}}=\frac{1}{\lambda}R_{\frac{1}{\lambda}}U. \tag{6}$$

关系式(4)和(6) 使我们可以将对 B_λ 所做的全部命题转述到算子 R_μ 上去,反之亦然.

4.2. 我们来研究当 λ 很小时豫解式 B_λ 的性态. 考察如下的级数

$$I+\lambda U+\lambda^2U^2+\cdots+\lambda^nU^n+\cdots \tag{7}$$

如果此级数在算子空间 $B(\mathbf{X},\mathbf{X})$ 中收敛, 则由 Banach 定理(见V. 4.5)的注可知其和为 $(I-\lambda U)^{-1}$, 亦即

$$(I-\lambda U)^{-1}=I+\lambda U+\cdots+\lambda^nU^n+\cdots.$$

因此, 根据(5)得

$$B_\lambda=U+\lambda U^2+\cdots+\lambda^nU^{n+1}+\cdots, \tag{8}$$

这个公式对于使得级数(7)收敛的 λ 均成立. 但正如在 V.4.2 中所证明的, 如果

$$\lim_{n\to\infty}\sqrt[n]{\|\lambda^nU^n\|}<1,$$

则级数(7)收敛; 如果

$$\lim_{n\to\infty}\sqrt[n]{\|\lambda^nU^n\|}>1,$$

则级数(7)发散. 因此我们得到下面的定理.

定理 1. 豫解式 B_λ 可展开为 λ 的幂级数(8), 其收敛半径为

$$r=\frac{1}{\lim\limits_{n\to\infty}\sqrt[n]{\|U^n\|}}=\frac{1}{\lim\limits_{n\to\infty}\sqrt[n]{\|B_0^n\|}}.$$

如果用(4)将 B_λ 变换为 R_μ, 则得下面推论.

推论. 豫解式 R_μ 可展开为 μ^{-1} 的幂级数

$$R_\mu=\frac{1}{\mu}I+\frac{1}{\mu^2}U+\cdots+\frac{1}{\mu^n}U^{n-1}+\cdots$$

$$\left(|\mu|>\frac{1}{r}=\lim_{n\to\infty}\sqrt[n]{\|U^n\|}\right).$$

4.3. 可以指出级数(8)收敛半径的另一表达式, 此式与特征集合在复平面上的分布有关.

先来证明两个引理.

引理 1. 对任何 $\lambda,\mu\in\pi(U)$, 等式

$$B_\lambda - B_\mu = (\lambda-\mu)B_\mu B_\lambda \tag{9}$$

成立.

证. 从关系式(5)可得

$$B_\lambda - B_\mu = U(I-\lambda U)^{-1} - (I-\mu U)^{-1}U.$$

先以$(I-\lambda U)$右乘，再以$(I-\mu U)$左乘上式，我们得到

$$\begin{aligned}(I-\mu U)(B_\lambda - B_\mu)(I-\lambda U) &= (I-\mu U)U - U(I-\lambda U)\\ &= (\lambda-\mu)U^2,\end{aligned}$$

因而

$$B_\lambda - B_\mu = (\lambda-\mu)(I-\mu U)^{-1}U\cdot U(I-\lambda U)^{-1} = (\lambda-\mu)B_\mu B_\lambda,$$

这就是所要证明的结果.

推论. 算子 B_λ 和B_μ是可交换的，即 $B_\lambda B_\mu = B_\mu B_\lambda$.

可类似地证明，对于任何 $\lambda, \mu\in\rho(U)$，成立等式

$$R_\lambda - R_\mu = -(\lambda-\mu)R_\lambda R_\mu.$$

引理 2. 像解式 B_λ 在集合 $\pi(U)$ 的每一点是参数 λ 的连续函数，也就是说：如果 $\lambda_n \to \lambda_0(\lambda_n, \lambda_0\in\pi(U))$，则 $B_{\lambda_n} \to B_{\lambda_0}$.

证. 首先来证明实函数$\|B_\lambda\|$在 $\pi(U)$ 上连续. 如果 $U=\mathbf{0}$，则 $B_\lambda=\mathbf{0}$，于是命题得证. 而如果 $U\neq\mathbf{0}$，则 $B_\lambda\neq\mathbf{0}$，从而可去证明函数$\frac{1}{\|B_\lambda\|}$的连续性. 从式(9)得

$$\begin{aligned}|\,\|B_\lambda\| - \|B_\mu\|\,| &\leqslant \|B_\lambda - B_\mu\| = |\lambda-\mu|\,\|B_\mu B_\lambda\|\\ &\leqslant |\lambda-\mu|\,\|B_\mu\|\,\|B_\lambda\|.\end{aligned}$$

因而

$$\left|\frac{1}{\|B_\mu\|} - \frac{1}{\|B_\lambda\|}\right| \leqslant |\lambda-\mu|,$$

由此即得所需要的结果.

现在来建立 B_λ 的连续性. 因 $\pi(U)$是开集而 $\lambda_0\in\pi(U)$，所以可找到一个完全包含在 $\pi(U)$之中的圆 $|\lambda-\lambda_0|\leqslant\varepsilon$. 连续函数$\|B_\lambda\|$在此圆内有界，例如可设

$$\|B_\lambda\|\leqslant M \qquad (|\lambda-\lambda_0|\leqslant\varepsilon). \tag{10}$$

由(9)和(10)得到

$$\|B_\lambda - B_{\lambda_0}\| \leqslant |\lambda-\lambda_0|\,\|B_{\lambda_0}\|\,\|B_\lambda\|$$
$$\leqslant M^2|\lambda-\lambda_0| \qquad (|\lambda-\lambda_0|\leqslant\varepsilon).$$

引理得证.

定理 2. 级数 (8) 的收敛半径 r 等于从点 $\lambda=0$ 到特征集合 $\chi(U)$的距离 r_0.

证. 首先,因在圆$|\lambda|<r$ 中级数(8)收敛. 因而对此圆内的 λ 豫解式存在,故此圆包含在非奇异值的集合之中. 所以 $r\leqslant r_0$.

现取任意的元素 $x\in\mathbf{X}$ 及任意的泛函 $f\in\mathbf{X}^*$ 并考察复变量 λ的函数 φ:

$$\varphi(\lambda)=f(B_\lambda(x)).$$

我们来证明 φ 在集合 $\pi(U)$上是正则的. 事实上, 如果 $\lambda,\mu\in\pi(U)$, 则由(9)

$$\frac{\varphi(\mu)-\varphi(\lambda)}{\mu-\lambda}=\frac{f(B_\mu(x))-f(B_\lambda(x))}{\mu-\lambda}=f(B_\mu B_\lambda(x)).$$

当 $\mu\to\lambda$ 时上式右端有极限 $f(B_\lambda^2(x))$ (见引理 2). 因此存在连续导数

$$\varphi'(\lambda)=f(B_\lambda^2(x)).$$

将函数 φ 在点 $\lambda_0=0$ 的邻域中展为 Taylor 级数

$$\varphi(\lambda)=\varphi(0)+\frac{\varphi'(0)}{1!}\lambda+\cdots+\frac{\varphi^{(n)}(0)}{n!}\lambda^n+\cdots. \tag{11}$$

此展式在不包含函数 φ 的奇点的圆内成立. 故在圆$|\lambda|<r_0$ 内成立. 但由(8)式

$$\varphi(\lambda)=\sum_{n=0}^{\infty}f(U^{n+1}(x))\lambda^n \qquad (|\lambda|<r). \tag{12}$$

这时由复变函数论中大家所熟悉的定理知级数(11)和(12) 恒等, 从而级数(12)当$|\lambda|<r_0$ 时收敛.

任取 $0<\lambda_1<r_0$, 从级数(12)在 $\lambda=\lambda_1$ 时收敛推出

$$\lambda_1^n f(U^{n+1}(x)) \xrightarrow[n\to\infty]{} 0,$$

因而，由 f 的任意性可知

$$\lambda_1^n U^{n+1}(x) \to \mathbf{0} \quad (\sigma(\mathbf{X}, \mathbf{X}^*)).$$

但弱收敛的元素序列是有界的(VIII. 1. 1):

$$\sup_n \|\lambda_1^n U^{n+1}(x)\| < \infty.$$

因为这个不等式对于每个 $x \in \mathbf{X}$ 均满足，而空间 $\mathbf{X}$ 完备，所以由定理 VII. 1. 1 得知

$$\sup_n \|\lambda_1^n U^{n+1}\| = M < \infty. \tag{13}$$

因此

$$\lambda_1 \lim_{n\to\infty} \sqrt[n]{\|U^n\|} \leqslant \lim_{n\to\infty} M^{\frac{1}{n}} = 1,$$

即

$$\lambda_1 \leqslant \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\|U^n\|}} = r.$$

因可取 λ_1 与 r_0 任意接近，所以 $r_0 \leqslant r$. 再由前面已经得到的不等式 $r \leqslant r_0$，故得 $r = r_0$. 证毕.

注 1. 设 λ_0 是算子 U 的非奇异值，和上面一样，展开式

$$B_\lambda = B_{\lambda_0} + (\lambda - \lambda_0) B_{\lambda_0}^2 + \cdots + (\lambda - \lambda_0)^n B_{\lambda_0}^{n+1} + \cdots$$

在圆 $|\lambda - \lambda_0| < \rho_0$ 中成立，其中 ρ_0 是从点 λ_0 到特征集合的距离，或如定理 1 所述

$$\rho_0 = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\|B_{\lambda_0}^{n+1}\|}}.$$

将豫解式 B_λ 换为豫解式 B_μ 则我们就得到如下结果.

推论 1. 展开式

$$R_\mu = \frac{1}{\mu} I + \frac{1}{\mu^2} U + \cdots + \frac{1}{\mu^{n+1}} U^n + \cdots \tag{14}$$

当 $|\mu|>\frac{1}{r}$ 时成立，其中 $\frac{1}{r}$ 是中心在坐标原点而完全包含全部谱点的最小圆的半径.

数 $1/r$ 叫作算子 U 的谱半径.

注 2. 如果 U 是 Hilbert 空间中的自共轭算子，则由IX. 5. 3可知 $\frac{1}{r}=\|U\|$. 将这和定理 1 推论的结果加以对照，即得有趣的关系式

$$\|U\|=\lim_{n\to\infty}\sqrt[n]{\|U^n\|}\ ^{*)}.$$

推论 2. 复 B-空间中连续线性算子 U 的谱 $\sigma(U)$ 是非空集合.

证. 如果 $\sigma(U)=\varnothing$，注意到 $\sigma(U)$ 和 $\chi(U)$ 的关系以及 $0\in\chi(U)$，我们就得到：非奇异值的集合 $\pi(U)$ 是整个复平面. 因可以认为 $U\neq\mathbf{0}$，所以对任何 λ，算子 $B_\lambda\neq\mathbf{0}$. 设 $y_0=B_{\lambda_0}(x_0)\neq\mathbf{0}$. 取 $f\in\mathbf{X}^*, f(y_0)\neq 0$. 和定理 2 的证明相似，我们得到函数

$$\varphi(\lambda)=f(B_\lambda(x_0))$$

在整个复平面上正则. 其次，还有估计式

$$|\varphi(\lambda)|=|f(B_\lambda(x_0))|\leqslant\|f\|\|B_\lambda\|\|x_0\|.$$

因为 $\sigma(U)=\varnothing$，故存在连续的 U^{-1}，因此，和引理 2 类似，我们就得知当 $\lambda\to\infty$ 时 $\|R_{1/\lambda}\|\to\|U^{-1}\|$. 因而，由(6)可得

$$\|B_\lambda\|\leqslant 1/|\lambda|\,\|R_{1/\lambda}\|\,\|U\|\underset{\lambda\to\infty}{\longrightarrow}0.$$

这说明 $\varphi(\lambda)$ 有界. 由 Liouville 定理可知 $\varphi(\lambda)$ 是恒常的(显然，只能等于 0). 但 $\varphi(\lambda_0)=f(y_0)\neq 0$，矛盾.

---

*) 但这个等式还只是一个更强的关系式

$$\|U\|=\sqrt[n]{\|U^n\|}\qquad(n=1,2,\cdots)$$

的推论. 要证明上述关系式只须将关于算子 $\varphi(U)$ 的谱定理(定理 IX. 5. 4)应用于函数 $\varphi(t)=t^n(n=1,2,\cdots)$.

需要在此指出：如果我们想用在 3.1 中那样的办法来确定实空间中算子 U 的谱，则谱可能是空集．这正是我们要在复空间考察问题的主要理由．

4.4. 我们利用所证得的结果来研究对方程

$$x-U(x)=y \tag{15}$$

应用逐次逼近法时的收敛性问题．

正象在 V.5.1 中所指出的那样，级数

$$I+U+U^2+\cdots+U^n+\cdots \tag{16}$$

的收敛性保证了以任意的 x_0 为初值的逐次逼近法的收敛性．在(14)中令 $\mu=1$，我们就得到如下的对方程(15)的逐次逼近法的收敛性准则．

定理 3． 如果算子 U 的谱位于圆 $|\mu|<1$ 内，则方程(15)对以任意 $x_0\in\mathbf{X}$ 为初值和任何右端 $y\in\mathbf{X}$ 的逐次逼近法均收敛．如果在圆 $|\mu|\leqslant 1$ 外有谱点，则存在集合*) $E\subset\mathbf{X}$ 使当 $y\in E$ 时以 $x_0=0$ 为初值对(15)的逐次逼近过程发散．

只需证明定理的第二部分．注意，当圆 $|\mu|\leqslant 1$ 外有谱点时，

$$\lim_{n\to\infty}\|U^n\|=\infty,$$

因此，由定理 VII.1.1 的注可知，除去 $\mathbf{X}$ 中某个第一纲集 G 外，对所有的 $y\in\mathbf{X}$ 均有

$$\sup_n\|U^n(y)\|=\infty. \tag{17}$$

但是，以 $x_0=0$ 为初值的逐次逼近过程的收敛性等价于级数

$$y+U(y)+U^2(y)+\cdots+U^n(y)+\cdots$$

的收敛性，而由(17)，如果 $y\notin G$，这个级数发散．

注 1． 如果算子 U 的所有谱点都是非零特征值，则定理还可

---

*) 集合 E 叫空间 $\mathbf{X}$ 中的留集．所谓留集乃是指 E 的余集是 $\mathbf{X}$ 中的第一纲集(见 I.4.7)．

表述成更确定的形式：逐次逼近法在这种情形收敛的充分必要条件是算子U的所有特征值均位于圆$|\mu|<1$内.

事实上，由上面的定理，如果逐次逼近法收敛，则算子U的谱位于圆$|\mu|\leqslant 1$内．如果假定在圆周$|\mu|=1$上存在特征值μ_0，在方程（15）中令$y=y'$，其中y'是对应于特征值μ_0的特征元，并取$x_0=\mathbf{0}$，我们得到n阶逼近x_n的表达式

$$x_n=(1+\mu_0+\cdots+\mu_0^{n-1})y',$$

它没有极限.

注 2. 如果算子U是 Hilbert 空间中的自共轭算子，则定理和注 1 将大可简化．由 4.3 中的注 2，在这种情形，我们就有：对方程(15)的逐次逼近法收敛的充分条件是$\|U\|<1$．如果U的所有非零谱点都是特征值，则此条件也是必要的.

最后，我们用特征集合的术语来叙述定理 3.

定理 3′. 如果算子U的特征集合位于圆$|\lambda|\leqslant 1$外，则方程(15)对任何右端$y\in\mathbf{X}$和以任意$x_0\in\mathbf{X}$为初值的逐次逼近法均收敛．如果在圆$|\lambda|<1$内有特征集合的点，则存在集合$E\subset\mathbf{X}$使当$y\in E$时以$x_0=\mathbf{0}$为初值的逐次逼近过程发散．这里的集合E为$\mathbf{X}$中的留集.

对于此定理，同样可给出类似于注 1 和注 2 的注，请读者自己来完成。

4.5. 如果U是紧算子，将定理 1 和定理 2 合并起来，即可得到关于豫解式在特征值邻域中的性质的更精细的事实.

如果连续线性算子V映空间$\mathbf{X}$到一有限维子空间$\widetilde{\mathbf{X}}\subset\mathbf{X}$内，则称算子$V$是有限维算子．取$\widetilde{\mathbf{X}}$中完全的线性独立的元素组$x_1, x_2, \cdots, x_n$．根据完全组的定义，对任何$x\in\mathbf{X}$

$$V(x)=\sum_{k=1}^{n}\alpha_k x_k,$$

其中系数$\alpha_1, \alpha_2, \cdots, \alpha_n$显然依赖于$x$．令$\alpha_k=f_k(x)$ $(k=1, 2, \cdots, n)$，不难验证泛函f_k $(k=1, 2, \cdots, n)$是线性的与连续的．其线性性是毫无疑问的，而其连续性由如下命题(见 IV. 1. 6)可以推知：

如果有限维赋范空间中的元素序列收敛于零，则其每个坐标也收敛于零．由此可得

$$V(x)=f_1(x)x_1+f_2(x)x_2+\cdots+f_n(x)x_n. \tag{18}$$

反之，所有形如(18)的算子 V 当然是有限维的.

还要指出一点：有限维算子必然是紧的.

设 U 是紧算子而 λ_0 是其特征值，则如下定理成立.

定理 4. 算子 U 可以表示为

$$U=U'+U''$$

的形式，其中 U'' 是有限维算子而 U' 是紧算子，并且，算子 U'' 的特征集合由唯一的一个点 λ_0 构成，算子 U' 的特征集合由算子 U 的特征集合除去点 λ_0 而构成.

证. 可以认为 $\lambda_0=1$(不然的话，代替 U 可去考察算子 $(1/\lambda_0)U$).在此假定下，我们来证明：将算子 U 分解为满足此定理条件的，定理 1.1 中所示的和 $U'+U''$.

沿用在定理 1.1 中所使用的符号，我们来验证：算子 U'' 有唯一的特征值 $\lambda_0=1$. 事实上，如果 $x_0\in\mathbf{N}(T)$，则更有 $x_0\in\mathbf{N}(T')=\mathbf{X}''$，因而，由算子 U'' 的定义

$$U''(x_0)=U(x_0)=x_0,$$

即 $\lambda_0=1$ 是算子 U'' 的特征值.

如果对某个 $x(x\in\mathbf{X}, x\neq\mathbf{0})$ 有

$$\lambda U''(x)=x, \tag{19}$$

则因 $U''(x)\in\mathbf{X}''$，所以 $x\in\mathbf{X}''$，因而 $U''(x)=U(x)$. 设 $m\geqslant 0$ 使得 $T^{m+1}(x)=\mathbf{0}$ 而 $T^m(x)\neq\mathbf{0}$. 由(19)

$$\begin{aligned}\mathbf{0}&=T^m(x-\lambda U(x))=T^m(x-\lambda[x-T(x)])\\&=(1-\lambda)T^m(x)+\lambda T^{m+1}(x)=(1-\lambda)T^m(x),\end{aligned}$$

可见，仅当 $\lambda=1$，上式才能成立.

于是，算子 U'' 的唯一特征值是 $\lambda_0=1$.

现在来证明定理的结论中关于算子 U' 的特征集合的部分.

因由定理 1.1 知算子 $T'=I-U'$ 有逆算子，故知 $\lambda_0=1$ 不是算子 U' 的特征值. 设 $\lambda\neq 1$ 是算子 U 的某个特征值而 $x_0\neq\mathbf{0}$ 是相应于 λ 的特征元. 如果 $x_0\in\mathbf{X}''$, 则如上面所证，我们就会得到 $\lambda=1$. 因此, $x_0\notin\mathbf{X}''$. 所以在展开式

$$x_0=x_0'+x_0'' \qquad (x_0'\in\mathbf{X}', x_0''\in\mathbf{X}'') \tag{20}$$

(参见定理 1.1 的 c)) 中应有 $x_0'\neq\mathbf{0}$. 根据展开式 (20) 的唯一性，从关系式

$$x_0=\lambda U(x_0')+\lambda U(x_0'')$$

得到 $x_0'=\lambda U(x_0')=\lambda U'(x_0')$, 即 λ 是算子 U' 的特征值.

反之，设 λ 是算子 U' 的特征值而 $x\neq\mathbf{0}$ 是相应的特征元. 因为

$$x=\lambda U'(x)\in\mathbf{X}',$$

所以 $U'(x)=U(x)$, 因而 $x=\lambda U(x)$, 即 λ 是算子 U 的特征值.

定理中其余的结论包含在定理 1.1 中.

定理证毕.

4.6. 关于豫解式在特征值邻域中的性态的更完全的表达可以在下述定理的基础上建立起来.

定理 5. 设 λ_0 是紧算子 U 的特征值，则在点 λ_0 的充分小的邻域内成立如下的展开式

$$B_\lambda=\frac{U_{-r}}{(\lambda-\lambda_0)^r}+\cdots+\frac{U_{-1}}{\lambda-\lambda_0}+U_0+U_1(\lambda-\lambda_0)+\cdots \\ +U_n(\lambda-\lambda_0)^n+\cdots \tag{21}$$

并且 r 是特征值 λ_0 的秩, 算子 $U_{-r},\cdots,U_{-1}$ 是有限维的, 算子

$$U_{-r}\neq\mathbf{0}.$$

(21)式右端的级数在算子空间 $B(\mathbf{X},\mathbf{X})$ 中收敛.

证. 如同在上一定理中那样，我们仍然认为 $\lambda_0=1$. 由 1.1 的

引理 2 立即可知特征值 λ_0 的秩是有限的. 再次利用定理 1.1 的表示并由这个定理的注, 可得 $\mathbf{X}'=T^r(\mathbf{X}), \mathbf{X}''=\mathbf{N}(T^r)$

将元素 $x\in\mathbf{X}$ 表示成如下形式

$$x=x'+x'' \qquad (x'\in\mathbf{X}', x''\in\mathbf{X}'')$$

(参见定理 1.1 的c)) 并将元素 x 和 $x'=P'(x)$ 或 $x''=P''(x)$ 对应起来, 我们就构造了从空间 $\mathbf{X}$ 到子空间 $\mathbf{X}'$ 和 $\mathbf{X}''$ 的两个投影算子 P' 和 P''. 由 §1 的估计式(9)可知这两个算子都是连续算子. 此外, 还有

$$P'(\mathbf{X})=\mathbf{X}', \quad P''(\mathbf{X})=\mathbf{X}'', \quad P'+P''=I.$$

考察任意元素 $y\in\mathbf{X}$. 元素 $x=B_\lambda(y)$ 是方程

$$x-\lambda U(x)=U(y) \tag{22}$$

的解. 将 $x=P'(x)+P''(x), y=P'(y)+P''(y)$ 代入上式并注意到 $U(\mathbf{X}')\subset\mathbf{X}'$ 和 $U(\mathbf{X}'')\subset\mathbf{X}''$, 我们可将(22)改写成如下的方程组

$$\left.\begin{array}{l} P'(x)-\lambda UP'(x)=UP'(y) \\ P''(x)-\lambda UP''(x)=UP''(y) \end{array}\right\}. \tag{23}$$

因为 $UP'=U'P'$, 故可把第一个方程改写为

$$T'_\lambda(P'(x))=U'P'(y),$$

其中 $T'_\lambda=I-\lambda U'$. 由于定理 4, $\lambda_0=1$ 是算子 U' 的正则值. 因此, 根据定理 2 的注, 如果$(\lambda-1)$充分小, 则算子 U' 的豫解式 B'_λ 可展为

$$B'_\lambda=U'_0+(\lambda-1)U'_1+\cdots+(\lambda-1)^nU'_n+\cdots,$$

且上式右端的级数在空间 $B(\mathbf{X}, \mathbf{X})$ 中收敛. 从上所述, 可得

$$P'(x)=B'_\lambda P'(y)=[U_0+(\lambda-1)U_1+\cdots+(\lambda-1)^nU_n+\cdots](y), \tag{24}$$

其中 $U_n=U'_nP'(n=1,2,\cdots)$ 并且上式右端的级数仍在空间 $B(\mathbf{X}, \mathbf{X})$ 中收敛.

现在来讨论(23)中第二个方程.

构造商空间

$$\mathbf{X}^{(k)}=\mathbf{N}(T^k)/\mathbf{N}(T^{k-1}) \qquad (k=1,2,\cdots,r),$$

并用 φ_k 来表示从空间 $\mathbf{N}(T^k)$ 到 $\mathbf{X}^{(k)}$ 上的自然同态. 空间 $\mathbf{X}^{(r)}$ 显然是有限维的. 取 $\mathbf{X}^{(r)}$ 中完全的线性独立元素组 $\bar{x}_1{}^{(r)}$, $\bar{x}_2{}^{(r)}$, $\cdots$, $\bar{x}_{n_r}^{(r)}$ 并设 $x_1{}^{(r)}, x_2{}^{(r)}, \cdots, x_{n_r}^{(r)}$ 是 $\mathbf{N}(T^r)$ 中使得 $\varphi_r(x_j{}^{(r)})=\bar{x}_j{}^{(r)}$ 的对应元素. 元素组 $x_j{}^{(r-1)}=T(x_j{}^{(r)})(j=1,2,\cdots,n_r)$ 属于 $\mathbf{N}(T^{r-1})$. 此外, 它们的象 $\bar{x}_j{}^{(r-1)}=\varphi_{r-1}(x_j{}^{(r-1)})$ 是线性独立的. 事实上, 若 $\sum_{j=1}^{n_r}\alpha_j\bar{x}_j{}^{(r-1)}=\mathbf{0}$, 则

$$\sum_{j=1}^{n_r}\alpha_j x_j{}^{(r-1)}=T\left(\sum_{j=1}^{n_r}\alpha_j x_j{}^{(r)}\right)\in\mathbf{N}(T^{r-2});$$

换句话说, $\sum_{j=1}^{n_r}\alpha_j x_j{}^{(r)}\in\mathbf{N}(T^{r-1})$, 因而

$$\sum_{j=1}^{n_r}\alpha_j\bar{x}_j{}^{(r)}=\varphi_r\left(\sum_{j=1}^{n_r}\alpha_j x_j{}^{(r)}\right)=\mathbf{0},$$

这仅当 $\alpha_1=\alpha_2=\cdots=\alpha_{n_r}=0$ 时才可能.

把元素 $\bar{x}_1{}^{(r-1)}, \bar{x}_2{}^{(r-1)}, \cdots, \bar{x}_{n_r}{}^{(r-1)}$ 和 $\bar{x}_{n_r+1}^{(r-1)}, \cdots, \bar{x}_{n_{r-1}}^{(r-1)}\in\mathbf{X}^{(r-1)}$ 合并起来使之构成 $\mathbf{X}^{(r-1)}$ 中的基. 其次, 选择 $x_{n_r+1}^{(r-1)}, \cdots, x_{n_{r-1}}^{(r-1)}\in\mathbf{N}(T^{r-1})$ 使得 $\bar{x}_j{}^{(r-1)}=\varphi_{r-1}(x_j{}^{(r-1)})$.

继续上述的讨论, 对于每个 $k=1,2,\cdots,r$ 构造元素 $x_1{}^{(k)}$, $x_2{}^{(k)}, \cdots, x_{n_k}{}^{(k)}$ 使得

$$x_j{}^{(k)}\in\mathbf{N}(T^k),\ T(x_j{}^{(k)})=x_j{}^{(k-1)} \tag{25}$$

$$(j=1,2,\cdots,n_k;\ k=1,2,\cdots,r;\ x_j{}^{(0)}=\mathbf{0}).$$

并且, 对于 $k=1,2,\cdots,r$, 元素

$\bar{x}_1{}^{(k)}=\varphi_k(x_1{}^{(k)}), \bar{x}_2{}^{(k)}=\varphi_k(x_2{}^{(k)}), \cdots, \bar{x}_{n_k}^{(k)}=\varphi_k(x_{n_k}^{(k)})$ 构成空间 $\mathbf{X}^{(k)}$ 的基.

命

$$\mathbf{X}_k=\mathscr{L}(\{x_1^{(k)}, x_2^{(k)}, \cdots, x_{n_k}^{(k)}\}) \qquad (k=1, 2, \cdots, r).$$

由(25)可知

$$T(\mathbf{X}_k)\subset\mathbf{X}_{k-1} \qquad (k=1, 2, \cdots, r; \mathbf{X}_0=\{\mathbf{0}\}). \tag{26}$$

今证元素组$\{x_j^{(k)}\}(j=1, 2, \cdots, n_k; k=1, 2, \cdots, r)$是$\mathbf{X}''$中完全的线性独立元素组.

设

$$\sum_{k=1}^{r}\sum_{j=1}^{n_k}\lambda_j^{(k)}x_j^{(k)}=\mathbf{0}.$$

因为当$k<r$时$x_j^{(k)}\in\mathbf{N}(T^{r-1})$，以算子$\varphi_r$作用上式，则

$$\sum_{k=1}^{r}\sum_{j=1}^{n_k}\lambda_j^{(k)}\varphi_r(x_j^{(k)})=\sum_{j=1}^{n_r}\lambda_j^{(r)}\bar{x}_j^{(r)}=\mathbf{0},$$

因而$\lambda_j^r=0(j=1, 2, \cdots, n_r)$. 用类似的办法可证所有其他的系数$\lambda_j^{(k)}$均等于零. 现在来考察任意元素$x''\in\mathbf{X}''$. 元素$\varphi_r(x'')\in\mathbf{X}^{(r)}$，因此，可找到系数$\alpha_1^{(r)}, \alpha_2^{(r)}, \cdots, \alpha_{n_r}^{(r)}$使得

$$\varphi_r(x'')=\sum_{j=1}^{n_r}\alpha_j^{(r)}\bar{x}_j^{(r)}=\varphi_r\Big(\sum_{j=1}^{n_r}\alpha_j^{(r)}x_j^{(r)}\Big).$$

所以

$$x''-\sum_{j=1}^{n_r}\alpha_j^{(r)}x_j^{(r)}\in\mathbf{N}(T^{r-1}).$$

继续如上的讨论，最后，我们可得到：对每个$\alpha_j^{(k)}$

$$x''-\sum_{k=1}^{r}\sum_{j=1}^{n_k}\alpha_j^{(k)}x_j^{(k)}\in\mathbf{N}(T^0)=\{\mathbf{0}\},$$

因而 $\quad x''=\sum_{k=1}^{r}\sum_{j=1}^{n_r}\alpha_j^{(k)}x_j^{(k)}.$

设x是$\mathbf{X}$中任意元素. 元素$P''(x)\in\mathbf{X}''$，因此

$$P''(x)=\sum_{k=1}^{r}\sum_{j=1}^{n_k}\alpha_j^{(k)}(x)x_j^{(k)},$$

并且，如同在 4.5 中指出的那样，系数 $\alpha_j^{(k)}$ 是线性泛函．如果记

$$P_k(x)=\sum_{j=1}^{n_k}\alpha_j^{(k)}(x)x_j^{(k)}\quad(k=1,2,\cdots,r),$$

根据上面的讨论，P_k 是映空间 $\mathbf{X}$ 到 $\mathbf{X}_k$ 的连续线性算子，并且

$$P''=P_1+P_2+\cdots+P_r,\tag{27}$$

$$P_mP_k=\begin{cases}\mathbf{0}, & m\neq k,\\ P_m, & m=k\end{cases}\quad(k,m=1,2,\cdots,r).\tag{28}$$

由于 $UP''=P''U$，将方程(23)中第二个的右端表示为 $P''U(y)$，然后用和 $P_1+P_2+\cdots+P_r$ 代替算子 P''：

$$\sum_{k=1}^{r}P_k(x)-\lambda\sum_{k=1}^{r}UP_k(x)=\sum_{k=1}^{r}P_kU(y).$$

以算子 P_m 作用上式两端并注意到(28)，则

$$P_m(x)-\lambda\sum_{k=1}^{r}P_mUP_k(x)=P_mU(y)\quad(m=1,2,\cdots,r).\tag{29}$$

但是

$$UP_k(x)=P_k(x)-TP_k(x),$$

所以，由(26)得

$$P_mUP_k(x)=\begin{cases}P_m(x), & k=m,\\ -TP_{m+1}(x), & k=m+1,\\ \mathbf{0}, & k\neq m,m+1,\end{cases}$$

$$(k,m=1,2,\cdots,r;P_{r+1}=\mathbf{0}).$$

利用这个关系式可将(29)写成更简单的形式

$$P_r(x)-\lambda P_r(x)=P_rU(y),\tag{30}$$

$$P_m(x)-\lambda[P_m(x)-TP_{m+1}(x)]=P_mU(y)\quad (m=1, 2, \cdots, r-1).\tag{31}$$

从(30)可得

$$P_r(x)=\frac{P_rU(y)}{1-\lambda}.$$

因而，根据(31)

$$\begin{aligned}P_{r-1}(x)&=\frac{\lambda}{1-\lambda}TP_r(x)+\frac{1}{1-\lambda}P_{r-1}U(y)\\&=\frac{\lambda}{(1-\lambda)^2}TP_rU(y)+\frac{1}{1-\lambda}P_{r-1}U(y).\end{aligned}$$

一般地，对任何 $m=1, 2, \cdots, r$

$$P_m(x)=\sum_{k=0}^{r-m}\frac{\lambda^k}{(1-\lambda)^{k+1}}T^kP_{m+k}U(y).$$

将所得到的等式加起来，得

$$\begin{aligned}P''(x)&=\sum_{m=1}^{r}P_m(x)=\sum_{m=1}^{r}\sum_{k=0}^{r-m}\frac{\lambda^k}{(1-\lambda)^{k+1}}T^kP_{m+k}U(y)\\&=\sum_{k=0}^{r-1}\frac{\lambda^k}{(1-\lambda)^{k+1}}\sum_{m=1}^{r-k}T^kP_{m+k}U(y).\end{aligned}\tag{32}$$

将分式$\dfrac{\lambda^k}{(1-\lambda)^{k+1}}$展为部分分式

$$\frac{\lambda^k}{(1-\lambda)^{k+1}}=\sum_{j=0}^{k}\frac{c_j^{(k)}}{(\lambda-1)^{j+1}},$$

这里的 $c_j^{(k)}$ 是某些常数，并且 $c_k^{(k)}=(-1)^{k+1}$. 将这些关系式代入(32)，得到

$$P''(x)=\sum_{s=1}^{r}\frac{U_{-s}(y)}{(\lambda-1)^s},\tag{33}$$

其中算子 $U_{-s}(s=1, 2, \cdots, r)$ 是形如 $T^kP_{m+k}U$ 的线性组合. 因为

$P_{m+k}(\mathbf{X})=\mathbf{X}_{m+k}\subset\mathbf{X}''$，由定理 1.1 的 b)，$T(\mathbf{X}'')\subset\mathbf{X}''$，所以算子 U_{-s} 是从空间 $\mathbf{X}$ 到 $\mathbf{X}''$的有限维算子．其次，由(32)可见

$$U_{-r}=(-1)^r T^{r-1}P_r U.$$

因此，譬如说，若 $y=x_1^{(r)}$，则由(25)

$$\begin{aligned}U_{-r}(y)&=(-1)^r T^{r-1}P_r U(x_1^{(r)})=(-1)^r T^{r-1}P_r(x_1^{r}-x_1^{(r-1)})\\&=(-1)^r T^{r-1}(x_1^{(r)})=(-1)^r x_1^{(1)}\neq\mathbf{0},\end{aligned}$$

于是 $U_{-r}\neq\mathbf{0}$.

把关系式(24)和(33)相加即得所要的豫解式 B_λ 的展开式．

定理全部证完．

注．如果 U 是 Hilbert 空间的自共轭算子，则上述定理还可以有某些改进．因为这时 $r=1$(见 3.1)，因而展开式(21)中只有一项 $(\lambda-\lambda_0)^{-1}U_{-1}$．关于这方面的结果的详细叙述此处就不再引进了，因为它在 IX.4.5 中已以更强的形式描述过．

§5. Fredholm 择一律

在这一节里，我们将指出将 B-空间映射为自身的连续线性算子 T 所应满足的一些条件，在这些条件下对算子 T 的 Fredholm 择一律成立(参见 1.4)．特别要指明的是，如果线性算子 U 的某次幂算子 U^m 是紧的，则对于算子 $T=I-U$ Fredholm 择一律成立．下面结果的叙述是 С. М. Никольский[1]给出的．

5.1. 考察方程

$$T(x)=y \tag{1}$$

及其共轭方程

$$T^*(g)=f. \tag{2}$$

另外，还考察相应的齐次方程

$$T(x)=\mathbf{0}, \tag{3}$$

$$T^*(g)=\mathbf{0}. \tag{4}$$

所谓对于算子T的 Fredholm 择一律成立乃是指:

1) 或者方程(1)和(2)对任何右端可解，这时其解是唯一的;

2) 或者是齐次方程(3)和(4) 分别具有个数相同的线性独立解$x_1, x_2, \cdots, x_n$和$g_1, g_2, \cdots, g_n$; 这时方程(1)或方程(2)可解的充分必要条件分别是

$$g_k(y)=0 \qquad (k=1,2,\cdots,n),$$

或

$$f(x_k)=0 \qquad (k=1,2,\cdots,n),$$

并且方程(1)和(2)的通解分别为

$$x=x^*+\sum_{k=1}^{n} c_k x_k$$

和

$$g=g^*+\sum_{k=1}^{n} c_k g_k,$$

其中x^*和g^*分别是方程(1)和(2)的某个特解,而$c_1, c_2, \cdots, c_n$是任意常数.

下面的定理表明使 Fredholm 择一律成立的算子T的类与形如$T=I-U$的算子类实质上区别不大,这里的U是紧算子.

定理 1. 下面条件中的每一个都是使 Fredholm 择一律对算子T成立的充分必要条件:

1) 算子T可表示为

$$T=W+V,$$

其中W是具有双边连续逆的算子,而V是紧算子.

2) 算子T可表示为

$$T=W_1+V_1,$$

其中W_1是具有双边连续逆的算子,而V_1是有限维算子.

证. 易见只要证明条件 1)的充分性和条件 2) 的必要性就可

以了.

条件 1 的充分性证明. 设

$$T=W+V$$

并且W有双边连续逆 W^{-1} 而 V 是紧算子. 此时方程(1)等价于

$$W^{-1}T(x)=W^{-1}(y). \tag{5}$$

其次,存在连续的双边逆算子 $W^{*-1}=(W^{-1})^{*}$(XII. 2. 3), 因此,在这样的意义下: 如果 g_0 是方程(6)的解, 则 $W^{*-1}(g_0)$是方程(2)的解;而如果 g_0' 是方程(2)的解,则 $W^{*}(g_0')$是方程(6)的解,方程(2)等价于

$$T^{*}W^{*-1}(g)=f. \tag{6}$$

引进记号 $U=-W^{-1}V$. 由于 $U^{*}=-V^{*}(W^{-1})^{*}=-V^{*}W^{*-1}$ (见 IX. 3. 1), 我们可把方程(5)和(6)改写为

$$x-U(x)=W^{-1}(y), \tag{7}$$

$$g-U^{*}(g)=f. \tag{8}$$

由于U是紧算子, 所以, 对方程(7)和(8)定理 1. 4 的结论成立. 因而齐次方程

$$x-U(x)=\mathbf{0}, \tag{9}$$

$$g-U^{*}(g)=\mathbf{0} \tag{10}$$

具有个数相同的线性独立解 $x_1, x_2, \cdots, x_n$ 及 $g_1', g_2', \cdots, g_n'$. 齐次方程(3)显然具有和方程(9)相同的完全线性独立解系,即 $x_1, x_2, \cdots, x_n$. 今证泛函

$$g_k=W^{*-1}(g_k') \qquad (k=1, 2, \cdots, n) \tag{11}$$

构成方程(4) 的完全线性独立解系. 由上面所指出的方程 (2)与(6)的等价性可推知(11)中每个泛函都是方程(4)的解. (11)中诸泛函是线性独立的,因为从关系式$\sum_{k=1}^{n}\alpha_k g_k=\mathbf{0}$ 推知$\sum_{k=1}^{n}\alpha_k g_k'=$

$\sum_{k=1}^{n}\alpha_k W^*(g_k)=\mathbf{0}$，而这只有在 $\alpha_1=\alpha_2=\cdots=\alpha_n=0$ 时才可能. 最后，如果方程(4)具有解 g_0，而 g_0 不是泛函(11)的线性组合，则泛函 $W^*(g_0)$ 将是方程(10)的不为泛函 $g_1', g_2', \cdots, g_n'$ 的线性组合的解，这是不可能的.

这就说明方程(3)和(4)具有个数相同的线性独立的解.

其次，仍由定理 1.4，方程(5)，因而也就是方程(1)有解的充分必要条件为

$$g_k'(W^{-1}(y))=0 \qquad (k=1,2,\cdots,n). \tag{12}$$

由于(11)，此条件与下述条件等价：

$$(W^{-1})^*(g_k')(y)=W^{*-1}(g_k')(y)=g_k(y)=0 \qquad (k=1,2,\cdots,n).$$

可类似地验证

$$f(x_k)=0 \qquad (k=1,2,\cdots,n)$$

是方程(2)可解的充分必要条件.

条件 2) 必要性的证明.

设 $x_1, x_2, \cdots, x_n$ 和 $g_1, g_2, \cdots, g_n$ 分别是方程(3)和(4)的完全的线性独立解系. 利用定理 V.7.4 和引理 III.3.1，可求得泛函 $f_1, f_2, \cdots, f_n\in\mathbf{X}^*$ 及元素 $y_1, y_2, \cdots, y_n\in\mathbf{X}$ 使得

$$f_j(x_k)=\begin{cases}0, & j\neq k,\\ 1, & j=k\end{cases} \qquad (j,k=1,2,\cdots,n), \tag{13}$$

$$g_k(y_j)=\begin{cases}0, & j\neq k\\ 1, & j=k\end{cases} \qquad (j,k=1,2,\cdots,n). \tag{14}$$

记 $\mathbf{Y}'=T(\mathbf{X})$，$\mathbf{Y}''=\mathscr{L}(\{y_1, y_2, \cdots, y_n\})$. 每个元素 $y\in\mathbf{X}$ 可唯一地表示成

$$y=y'+y'' \qquad (y'\in\mathbf{Y}', y''\in\mathbf{Y}''). \tag{15}$$

事实上，若设

$$y''=\sum_{k=1}^{n} g_k(y)y_k, \qquad y'=y-y'',$$

则由(14)

$$g_j(y')=g_j(y)-\sum_{k=1}^{n} g_k(y)g_j(y_k)=0 \quad (j=1,2,\cdots,n).$$

因此方程 $T(x)=y'$ 有解,从而 $y'\in\mathbf{Y}'$. 表示式(15)的唯一性可这样来推出: 如果 $y''=\sum_{k=1}^{n}\alpha_k y_k\in\mathbf{Y}'$, 则方程 $T(x)=y''$ 应存在解, 从而 $g_j(y'')=\alpha_j=0(j=1,2,\cdots,n)$.

其次, 记 $\mathbf{X}'=\mathbf{N}(f_1,f_2,\cdots,f_n)$, $\mathbf{X}''=\mathscr{L}(\{x_1,x_2,\cdots,x_n\})$. 类似于前面的证明,所有元素 $x\in\mathbf{X}$ 均可唯一地表示成

$$x=x'+x'' \qquad (x'\in\mathbf{X}',\ x''\in\mathbf{X}''). \tag{16}$$

构造算子 W_1 并设

$$W_1(x)=T(x)+\sum_{k=1}^{n} f_k(x)y_k,$$

现在来证明算子 W_1 是从空间 $\mathbf{X}$ 到自身的一对一的映射, 从而具有双边连续逆算子(定理 XII. 1. 2). 事实上, 设 y 是 $\mathbf{X}$ 中的任意元素, 将 y 表示成(15)的形式 $y=y'+y''$, 并且 $y'\in\mathbf{Y}'=T(\mathbf{X})$, $y''=\sum_{k=1}^{n}\alpha_k y_k\in\mathbf{Y}''$, 亦即方程 $T(x)=y'$ 有解 x', x' 可认为是 $\mathbf{X}'$ 中的元素 (事实上, 将解 x 表示如(16)式: $x=x'+x''$. 再由 $T(x'')=\mathbf{0}$, 则知 $y'=T(x)=T(x')$).

命

$$x''=\sum_{k=1}^{n}\alpha_k x_k, \quad x^*=x'+x'',$$

注意 $T(x'')=\mathbf{0}$, $f_j(x')=0$ 和(13), 即可知

$$W_1(x^*)=T(x')+T(x'')+\sum_{k=1}^{n}f_k(x'')y_k$$

$$=T(x')+\sum_{k=1}^{n}f_k(\sum_{j=1}^{n}\alpha_j x_j)y_k=y'+\sum_{k=1}^{n}\alpha_k y_k$$

$$=y.$$

现在来证明,除去元素 x^*外, 方程 $W_1(x)=y$ 不存在其他的解. 事实上,如果存在非零元素 x_0 使得

$$W_1(x_0)=\mathbf{0},$$

即

$$T(x_0)+\sum_{k=1}^{n}f_k(x_0)y_k=\mathbf{0}.$$

这时 $T(x_0)\in\mathbf{Y}'$ 而 $\sum_{k=1}^{n}f_k(x_0)y_k\in\mathbf{Y}''$. 由(15)式表示的唯一性,即得

$$T(x_0)=\mathbf{0},\quad \sum_{k=1}^{n}f_k(x_0)y_k=\mathbf{0}, f_k(x_0)=0\quad(k=1,2,\cdots,n),$$

由此得知 $x_0=\mathbf{0}$(因为同时有 $x_0\in\mathbf{X}''$ 及 $x_0\in\mathbf{X}'$).

为了完成定理的证明,只要令

$$V_1(x)=-\sum_{k=1}^{n}f_k(x)y_k$$

即可.

注. 建议读者证明: 如果在条件 1)或 2)中用算子 T^*代替算子 T, 则得到的两个条件也是对T成立 Fredholm 择一律的充分必要条件.

5.2. 更进一步的叙述要用到两个简单的引理.

引理 1. 设A和B是由赋范空间 $\mathbf{X}$ 到自身的连续线性算子. 如果这两个算子是可交换的并且 $C=AB$ 具有双边逆算子, 则算

子 A 和 B 同样具有逆算子.

证. 首先证明算子 A 和 C^{-1} 是可交换的. 事实上,

$$A=C^{-1}CA=C^{-1}ABA=C^{-1}A(AB)=C^{-1}AC.$$

上式两端右乘以 C^{-1} 即得 $AC^{-1}=C^{-1}A$.

其次, 利用已证的算子 A 与 C^{-1} 的可交换性可得

$$B(AC^{-1})=BAC^{-1}=CC^{-1}=I,$$

及

$$(AC^{-1})B=C^{-1}AB=C^{-1}C=I,$$

由此推出存在 $B^{-1}=AC^{-1}$. 可类似地证明存在 $A^{-1}=BC^{-1}$.

注. 如果算子 C^{-1} 是连续的, 则算子 A^{-1} 和 B^{-1} 也是连续的.

引理 2. 设 U 是空间 $\mathbf{X}$ 中的连续算子, 则算子 U 的特征集合 $\chi(U)$ 和算子 U^m 的特征集合 $\chi(U^m)$ 存在下述关系

$$[\chi(U)]^m \subset \chi(U^m),$$

即如果 $\lambda\in\chi(U)$, 则 $\lambda^m\in\chi(U^m)$.

证. 令 $\varepsilon=e^{2\pi i/m}$, 则

$$I-\lambda^m U^m=(I-\lambda U)(I-\lambda\varepsilon U)\cdots(I-\lambda\varepsilon^{m-1}U).$$

如果 $\lambda^m\notin\chi(U^m)$, 则令

$$A=I-\lambda U, B=(I-\lambda\varepsilon U)\cdots(I-\lambda\varepsilon^{m-1}U), C=I-\lambda^m U^m,$$

故存在连续逆算子 C^{-1}, 因而由引理 1 的注可知存在连续逆算子 A^{-1}, 即 $\lambda\notin\chi(U)$.

5.3. 如同在 5.1 中一样认为 $\mathbf{X}$ 是 B-空间, 我们来考察 $\mathbf{X}$ 中的连续线性算子 U.

定理2. 设存在使得 U^m 是紧算子的自然数 m, 则这时对算子 $T=I-U$ 的 Fredholm 择一律成立.

证. 由引理 2 知特征集合 $\chi(U)$ 由孤立点组成, 故在复平面的单位圆周上仅有有限个点 $\lambda_1, \lambda_2, \cdots, \lambda_\nu\in\chi(U)$.

设p遍历整个质数集合，数

$$e^{2k\pi i/p} \qquad (k=1,2,\cdots,p-1)$$

各不相同，因此对充分大的p_0

$$\lambda_j \neq e^{2k\pi i/p} \qquad (p\geqslant p_0; k=1,2,\cdots, p-1; j=1,2,\cdots,\nu). \tag{17}$$

可设$m\geqslant p_0$并且m是质数. 我们写出展开式

$$I-U^m=(I-U)(I-\varepsilon U)\cdots(I-\varepsilon^{m-1}U)=(I-U)V, \tag{18}$$

其中

$$\varepsilon=e^{2\pi i/m}, \quad V=(I-\varepsilon U)\cdots(I-\varepsilon^{m-1}U).$$

由(17)，算子$I-\varepsilon^k U(k=1,2,\cdots,m-1)$可逆，因此存在连续逆算子$V^{-1}$. 但这时

$$T=I-U=(I-U^m)V^{-1}=V^{-1}-U^mV^{-1}.$$

由于算子V^{-1}可逆，算子U^mV^{-1}是紧算子，故可以应用定理1.

定理证毕.

5.4. 关于紧算子特征值集合的定理(定理2)可推广到上一个定理所考察过的那种形式的算子上去. 即如下定理成立.

定理 3. 如果对某个m，算子U^m是紧的，则1)算子U的特征集合$\chi(U)$仅由特征值组成，并且每个特征值都具有有限秩，而对应的特征子空间是有限维的；2)在复平面的每个圆$|\lambda|\leqslant R$内仅有有限个特征值.

证. 根据引理2，可只限于定理中结论1)的证明. 另外，易见"算子U的特征集合$\chi(U)$仅由特征值组成"这一结论可由定理2推出. 因此只需证明每个特征值的秩都是有限的而且相应的特征子空间是有限维的.

不失一般性，可设所要考察的特征值是$\lambda_0=1$. 记$T_m=I-U^m$. 由(18)

$$T_m = VT, \quad T = V^{-1}T_m.$$

因为算子 T, T_m, V 和 V^{-1} 都是彼此可交换的，故

$$T_m^n = V^nT^n, \quad T^n = V^{-n}T_m^n \qquad (n=1, 2, \cdots).$$

由此立即推出

$$\mathbf{N}(T_m^n) = \mathbf{N}(T^n). \tag{19}$$

因此，由于 U^m 是紧算子，而对这种算子我们已经证明过命题是成立的，根据 (19) 可知对我们现在所考察的情形命题亦成立.

最后我们举一个例子，即连续线性算子 U 本身不是紧的，但算子 U^2 是紧的.

设 $\mathbf{X}$ 是空间 $l^p(1 \leqslant p \leqslant \infty)$, c, c_0 之一，对元素 $x = \{\xi_n\} \in \mathbf{X}$，命 $y = U(x)$，其中

$$y = \{\eta_n\} \quad \left(\eta_n = \begin{cases} 0, & n \text{ 为奇数} \\ \xi_{n-1}, & n \text{ 为偶数} \end{cases} \quad n = 1, 2, \cdots\right),$$

显然 $U^2 = \mathbf{0}$.

注. 由于 § 4 中对紧算子所证明的定理只用到定理 1.1 和定理 3.1 中包含的紧算子的性质，将这些定理转到上面讨论的那种形式的算子时情况并无什么变化，所以，若仅仅认为算子 U 的某次幂是紧的，则 § 4 中所得的结果对这样的算子也成立.

§ 6. 对积分方程的应用

6.1. 考察积分方程

$$x(s) - \lambda\int_0^1 K(s, t)x(t)dt = y(s), \tag{1}$$

假定核 $K(s, t)$ 是正方形 $[0, 1; 0, 1]$ 中的连续函数. 如果把积分项看成空间 $\mathbf{C}[0, 1]$ 中的线性算子，则方程(1)就是我们在上一节中所考察过的那种类型的方程.

还可以考察比方程(1)更为一般的积分方程，即

$$x(s)-\lambda\int_{T} K(s,t)x(t)dt=y(s), \tag{1'}$$

其中T是n维欧氏空间中的任意有界闭集(此时，s和t表示n维空间中的点)．但对方程(1)所作的全部证明不需作任何实质的改变即可搬到对方程(1′)的证明中来，因此，我们只讨论比较简单的方程(1)．

我们把积分算子

$$z=U(x),\quad z(s)=\int_0^1 K(s,t)x(t)dt \tag{2}$$

看作从$\mathbf{C}[0,1]$到$\mathbf{C}[0,1]$内的算子，其范数是

$$\|U\|=\max_s\int_0^1 |K(s,t)|dt$$

(参见V.2.4)，由IX.2.1可推U是紧算子．

根据U的定义，可将方程(1)写成

$$x-\lambda U(x)=y. \tag{3}$$

这个方程的解x^*通过y用公式

$$x^*=y+\lambda B_\lambda(y)$$

表示．由定理4.1，此解可展成幂级数

$$x^*=y+\lambda U(y)+\cdots+\lambda^n U^n(y)+\cdots, \tag{4}$$

这个级数对于满足不等式

$$|\lambda|<\frac{1}{d}=r$$

的所有的λ均收敛，其中$d=\lim\limits_{n\to\infty}\sqrt[n]{\|U^n\|}$，$r$是点$\lambda=0$到算子$U$的特征集合的距离(定理4.2)．级数(4)对于满足

$$|\lambda|<\frac{1}{\|U\|}=\frac{1}{\max\limits_s\int_0^1 |K(s,t)|dt}$$

的所有λ总是成立的．

正如在 V.5.5 中所指出的，算子 U 的幂也是积分算子，即

$$z=U^n(x), \qquad z(s)=\int_0^1 K_n(s,t)x(t)dt \qquad (n=1,2,\cdots), \tag{5}$$

其中 $K_n(s,t)$ 是重核.

将(5)代入(4)，即得到方程(1)按参数 λ 展开的解：

$$x^*(s)=y(s)+\lambda\int_0^1 K(s,t)y(t)dt+\cdots$$
$$+\lambda^n\int_0^1 K_n(s,t)y(t)dt+\cdots.$$

并且此级数对 $s\in[0,1]$ 一致收敛.

因为级数

$$B_\lambda=U+\lambda U^2+\cdots+\lambda^{n-1}U^n+\cdots \tag{6}$$

在从 $\mathbf{C}[0,1]$ 到 $\mathbf{C}[0,1]$ 的算子空间内收敛，则

$$\left\|\sum_{j=m+1}^{m+p}\lambda^{j-1}U^j\right\|=\max_s\int_0^1\left|\sum_{j=m+1}^{m+p}\lambda^{j-1}K_j(s,t)\right|dt\xrightarrow[m\to\infty]{}0.$$

因而，对每个固定的 $s\in[0,1]$，级数

$$\sum_{j=1}^{\infty}\lambda^{j-1}K_j(s,t) \tag{7}$$

在空间 $\mathbf{L}^1$ 中对 $s\in[0,1]$ 一致收敛. 记这个级数的和为函数 $\Gamma(s,t;\lambda)$，并叫作积分方程(1)的预解式. 显然

$$B_\lambda(y)(s)=\int_0^1\Gamma(s,t;\lambda)y(t)dt,$$

因此，公式(4)还可以改写为

$$x^*(s)=y(s)+\lambda\int_0^1\Gamma(s,t;\lambda)y(t)dt.$$

如果 $|\lambda|<r$，则由定理 4.3 知对方程(3)的逐次逼近过程收敛，把这个结论应用到积分方程(1)，我们就得到如下结果：

如果$|\lambda|<r$，则积分方程(1)的解可以作为按递推公式

$$x_{n+1}(s)=\lambda\int_0^1 K(s,t)x_n(t)dt+y(s) \qquad (n=0,1,2,\cdots)$$

所确定的连续函数序列$\{x_n(s)\}$的一致收敛的极限而得出，并且初值$x_0(t)$是任意的连续函数.

6.2. 如果核$K(s,t)$在$s<t$时为零，则方程(1)可以写成

$$x(s)-\lambda\int_0^s K(s,t)x(t)dt=y(s). \tag{8}$$

这种类型的方程叫做Volterra积分方程.

不难证明Volterra方程的重核当$s<t$时也为零.

假定核$K(s,t)$当$0\leqslant t\leqslant s\leqslant 1$时连续，今证展开式(4)对所有的复数$\lambda$都成立，即$r=\infty$.

闭集合上连续的函数必有界，故可认为$|K(s,t)|\leqslant M$. 对重核$K_n(s,t)$不难得到估计式

$$|K_n(s,t)|\leqslant\frac{s^{n-1}}{(n-1)!}M^n \qquad (n=1,2,\cdots). \tag{9}$$

事实上，$n=1$时(9)式显然成立，如果(9)式对某个$n>1$成立，则

$$|K_{n+1}(s,t)|\leqslant\int_0^s |K(s,u)K_n(u,t)|\,dt$$

$$\leqslant M\int_0^s\frac{u^{n-1}}{(n-1)!}M^n du=\frac{s^n}{n!}M^{n+1},$$

从(9)式得到

$$\|U^n\|=\max_s\int_0^s|K_n(s,t)|\,dt$$

$$\leqslant M^n\max_s\int_0^s\frac{s^{n-1}}{(n-1)!}dt=\frac{M^n}{(n-1)!},$$

由此

$$\sqrt[n]{\|U^n\|}\leqslant M\sqrt[n]{\frac{1}{(n-1)!}}\xrightarrow[n\to\infty]{}0.$$

这说明$r=\infty$，即Volterra型积分算子没有特征值.

6.3. 仍考察方程(1). 因为算子(2)是紧的，对于方程(3)来说 Fredholm 择一律成立. 由此得到对积分方程(1)的下述结果.

定理 1. 或者是积分方程(1)对于任意的连续函数 $y(s)$ 有唯一的连续解，或者方程

$$x(s)-\lambda\int_0^1 K(s,t)x(t)dt=0 \tag{10}$$

具有有限个线性独立的解 $x_1(s), x_2(s), \cdots, x_n(s)$. 这时方程

$$\psi(t)-\overline{\lambda}\int_0^1 \overline{K(s,t)}\psi(s)ds=0$$

具有和(10)的解数目相等的线性独立解 $\psi_1(t), \psi_2(t), \cdots, \psi_n(t)$. 当齐次方程(10)具有 n 个线性独立解时，方程(1)存在解的充分必要条件是

$$\int_0^1 \overline{\psi_k(s)}\, y(s)ds=0.$$

使(10)存在解的非零 λ 值叫方程(1)或核 $K(s,t)$ 的特征值. 换言之，方程(1)的特征值不是别的，正是算子 U 的特征值. 因此，对于积分方程的按模最小的特征值 λ_1 有估计式

$$|\lambda_1|=R=\frac{1}{d}\geqslant\frac{1}{\|U\|}$$

$$=\frac{1}{\max\limits_s\int_0^1|K(s,t)|dt}\geqslant\frac{1}{\max\limits_{s,t}|K(s,t)|}.$$

利用方程(3)的解依赖于 λ 的关系，由定理 4.5 得到以下结果.

定理 2. 在特征值 λ_0 的邻域中，方程(1)的解 x^* 可表示成

$$x^*(s)=\frac{y_{-r}(s)}{(\lambda-\lambda_0)^r}+\cdots+\frac{y_{-1}(s)}{\lambda-\lambda_0}+y_0(s)+\cdots$$

$$+y_n(s)(\lambda-\lambda_0)^n+\cdots,$$

其中 $y_k(s)(k=-r,\cdots,0,\cdots)$是仅依赖于 y 的连续函数，右端级数对 $s\in[0,1]$一致收敛.

因此，对于每一个固定的 $s\in[0,1]$，$x^*(s)$是在特征值处有极点的半纯函数.

6.4. 现在我们来考察空间 $\mathbf{L}^p(D)$（其中 D 是 n 维空间中的有界域）以及其中的积分方程

$$x(s)-\lambda\int_D K(s,t)x(t)dt=y(s)\quad(s\in D).\tag{11}$$

设核 $K(s,t)$满足定理 XI.3.3 的条件($q=p$)，即

1) $\left\{\int_D|K(s,t)|^r dt\right\}^{\frac{1}{r}}\leqslant C_1\qquad(s\in D);$

2) $\left\{\int_D|K(s,t)|^\sigma ds\right\}^{\frac{1}{\sigma}}\leqslant C_2\qquad(t\in D);$

3) $p-r(p-1)<\sigma<p.$

这时根据上面所引用的定理，以 $K(s,t)$ 为核的积分算子 U 是从 $\mathbf{L}^p(D)$到 $\mathbf{L}^p(D)$的紧算子，其范数满足估计式$\|U\|\leqslant C_1^{1-\sigma/p}C_2^{\sigma/p}$，因而对方程(11)可以运用上面对方程(1)所作的全部结论.

特别，如果核 $K(s,t)$是位势型的，即若

$$K(s,t)=B(s,t)/|s-t|^m,$$

其中 $B(s,t)$当 $s\neq t$ 时是有界连续函数，而 $m<n$，则上面所指的条件都成立(定理 XI.3.6).

此外，如果 $B(s,t)$是连续函数而

$$p>\frac{n}{n-m},$$

则由定理 XI.3.7，算子U将 $\mathbf{L}^p(D)$映射到 $\mathbf{C}(D)$，于是这时解 $x^*(s)$ 的幂级数展式不论在特征值邻域里还是在零点邻域里都是一致收敛的.

详细的叙述建议读者去做.

§7. 紧算子的不变子空间・逼近问题

本世纪60—70年代之间，在 Banach 空间的理论方面出现了一系列深刻的结果. 结果之一就是由 Lindenstrauss 和 Tzafriri 所给出的 Hilbert 空间的特征(见定理 V.3.3). 我们在这里还要建立两个重要的结果. (关于 Banach 空间理论的近代状况可见 Lindenstrauss 和 Tzafriri 的专著.)

7.1. 设 $\mathbf{X}$ 是 Banach 空间，U 是 $\mathbf{X}$ 中的连续线性算子. 闭线性集合 $E\subset\mathbf{X}$ 称为算子 U 的不变子空间，如果 $U(E)\subset E$ 的话. 易见$\{\mathbf{0}\}$和 $\mathbf{X}$ 均是不变子空间. 如果E既不为 $\mathbf{X}$ 又不为$\{\mathbf{0}\}$，则称E是非平凡的不变子空间. 特征向量算子的存在表明算子具有一维的不变子空间. 但是即使是紧算子也可能一个特征向量也没有.

Volterra 积分算子就是一例(见 6.2). 但这个算子却有很多不变子空间. 例如，对于任意的 $a\in(0,1)$，所有在区间$[0,a]$中等于零的函数 $x\in\mathbf{C}[0,1]$的集合就是这个算子的不变子空间.

J. von. Neumann 在1935年曾经证明：Hilbert 空间中任何非零紧算子均存在非平凡的不变子空间*). 不久前 В. И. Ломоносов[1] 得到了更为一般的事实的一个简单证明(我们在这里采用的是这个证明的一种变形，是由 Khilden 给出的.)

设U是无限维 Banach 空间 $\mathbf{X}$ 中的紧算子. 以 $\mathfrak{A}$ 表示所有与 U 可交换的算子的集合，即所有使得 $TU=UT$ 成立的算子 $T\in B(\mathbf{X},\mathbf{X})$的集合. 显然，$\mathfrak{A}$ 是$B(\mathbf{X},\mathbf{X})$中的线性子集合而且若有 T_1，$T_2\in\mathfrak{A}$，则 $T_1T_2\in\mathfrak{A}$.

---

*) J. von. Neumann 的这个结果见于 Aronszain 和 Smith 的文章[1]中，这里 Neumann 将之推广到 Banach 空间上.

定理 1 (Ломоносов). 在无限维 Banach 空间 $\mathbf{X}$ 中存在这样的子空间 E: E 是所有算子 $T\in\mathfrak{A}$ 的非平凡不变子空间.

证. 因为 $U\neq\mathbf{0}$, 故可认为 $\|U\|=1$. 另外, 还可以找到元素 $x_0\in\mathbf{X}$ 使得

$$\|x_0\|>1 \quad 而且 \quad \|U(x_0)\|>1. \tag{1}$$

设 B 是以 x_0 为中心的开的单位球. 因为算子 U 是紧的, 所以 $K=\overline{U(B)}$ 也是紧的. 任何 $x\in B$ 均可表示成 $x=x_0+z$, 其中 $\|z\|<1$. 如果 $y=U(x)$, 则 $y=U(x_0)+U(z)$ 并且 $\|U(z)\|\leqslant 1$. 因而

$$\|y\|\geqslant\|U(x_0)\|-\|U(z)\|\geqslant\|U(x_0)\|-1=\delta>0,$$

由此, 对任何 $y\in K$, 均有

$$\|y\|\geqslant\delta>0, \tag{2}$$

其中 δ 不依赖于 y. 我们来考察两种情况:

a) 设算子 U 具有非零特征值 λ. 用 N_λ 表示对应于 λ 的特征子空间. 我们来证明 N_λ 是任何 $T\in\mathfrak{A}$ 的不变子空间(因为 U 是紧的, 所以显然 $N_\lambda\neq\mathbf{X}$). 如果 $x\in N_\lambda$, 则

$$U(T(x))=T(U(x))=T(\lambda x)=\lambda T(x),$$

由此可知 $T(x)\in N_\lambda$.

b) 其次假定算子 U 没有特征值, 从而谱 $\sigma(U)=\{\mathbf{0}\}$. 因而算子 U 的谱半径等于零, 由此据 4.1 所述, 对任意的 λ

$$\|\lambda^n U^{n+1}\|\xrightarrow[n\to\infty]{}0. \tag{3}$$

假定不存在关于所有 $T\in\mathfrak{A}$ 的非平凡不变子空间. 对任意的 $x\in\mathbf{X}(x\neq\mathbf{0})$, 我们来考察子空间 $\mathbf{L}_x$, 这里的 $\mathbf{L}_x$ 等于集合 $\mathbf{L}'_x=\{T(x): T\in\mathfrak{A}\}$的闭包. 因为 $\mathfrak{A}$ 是线性集合, 所以集合 $\mathbf{L}'_x$ 也是线性的. 如果 $y\in\mathbf{L}'_x$, 则 $y=T_0(x)$, $T_0\in\mathfrak{A}$. 对于任何 $T\in\mathfrak{A}$, 均有 $TT_0\in\mathfrak{A}$, 因此可知 $T(y)=(TT_0)(y)\in\mathbf{L}'_x$. 由此并考虑到 $\mathfrak{A}$ 中所有算子的连续性, 则知 $\mathbf{L}_x$ 是 $\mathfrak{A}$ 中任何算子的不变子空间.

因为恒等算子 I 属于集合 $\mathfrak{A}$, 所以$x\in\mathbf{L}'_x$, 于是可知 $\mathbf{L}_x\neq\{\mathbf{0}\}$。

因为我们曾假定不存在非平凡的不变子空间，所以对任何 $x\neq\mathbf{0}$ 均有 $\mathbf{L}'_x=\mathbf{X}$，由此可知 $\mathbf{L}'_x\cap B$ 非空.

因而对任何 $x\neq\mathbf{0}$ 均存在 $T\in\mathfrak{A}$ 使得 $T(x)\in B$，由此可知

$$\bigcup_{T\in\mathfrak{A}}T^{-1}(B)=\mathbf{X}\setminus\{\mathbf{0}\}。$$

由(2)知 $K\subset\mathbf{X}\setminus\{\mathbf{0}\}$，由此可知$\{T^{-1}(B):\ T\in\mathfrak{A}\}$是 K 的紧开覆盖. 我们从中选取有限覆盖:

$$K\subset\bigcup_{i=1}^{m}T_i^{-1}(B).$$

取满足关系式(1)的点 x_0，因 $U(x_0)\in U(B)\subset K$，则可找到 $i_1(1\leqslant i_1\leqslant m)$使得 $U(x_0)\in T_{i_1}^{-1}(B)$. 因此

$$x_1=T_{i_1}(U(x_0))\in B.$$

类似地，可找到 $i_2(1\leqslant i_2\leqslant m)$使得

$$x_2=T_{i_2}(U(x_1))\in B,$$

如此等等. 这样就得到序列$\{x_n\}\subset B$，使得

$$x_n=T_{i_n}UT_{i_{n-1}}UT_{i_{n-2}}\cdots T_{i_1}U(x_0).$$

考虑到算子 T_i 与 U 是可交换的，则

$$x_n=T_{i_n}T_{i_{n-1}}T_{i_{n-2}}\cdots T_{i_1}U^n(x_0).$$

将算子 U 作用于 x_n，则

$$z_n=U(x_n)=T_{i_n}T_{i_{n-1}}T_{i_{n-2}}\cdots T_{i_1}U^{n+1}(x_0)\in K.$$

由(2)可知

$$\|z_n\|\geqslant\delta>0\qquad(n\in N).\tag{4}$$

现在我们来对$\|z_n\|$作另一估计. 设 $\lambda=\max\{\|T_i\|:\ 1\leqslant i\leqslant n\}$，则

$$\|z_n\|\leqslant\prod_{j=1}^{n}\|T_{i_j}\|\|U^{n+1}\|\|x_0\|\leqslant\|\lambda^nU^{n+1}\|\|x_0\|.\tag{5}$$

由(3)可知(5)式的右端当 $n\to\infty$ 时趋于零，这与(4)式矛盾.

这就证明了定理. 特别, 从所证的这个定理我们得知算子 U 本身具有非平凡的不变子空间, 此即 Neumann 的最初的结果.

正象我们业已看到的那样, 关于某类算子具有非平凡的不变子空间的问题是一个很不简单的问题. 而更为复杂的问题却是这个问题的逆问题: 求作用在可分 Banach 空间中的没有非平凡不变子空间的连续线性算子.

7.2. 正象我们在本章中已经讲过的那样, 紧算子就其本身的性质来说是最接近于有限维算子的. 这一事实促使 A. Grothendieck[2]给出下述定义:

我们说 Banach 空间 $\mathbf{X}$ 具有可逼近性质, 如果从任何 Banach 空间 $\mathbf{Y}$ 到 $\mathbf{X}$ 的任何紧算子 U 均是有限维算子序列的极限(按空间 $B(\mathbf{Y}, \mathbf{X})$ 的范数收敛于 U)的话. 这个条件等价于: 对任何紧的 $K \subset \mathbf{X}$ 和任何 $\varepsilon > 0$, 存在有限维算子 $T \in B(\mathbf{X}, \mathbf{X})$, 使得

$$\|Tx - x\| < \varepsilon \qquad (x \in K).$$

A. Grothendieck 提出这样的问题: 是否任何 Banach 空间均具有可逼近性质? 他证明了: 这个可逼近性问题的正面回答等价于满足一系列很具体的假定. 例如就可以是这样的: 如果 $K(s, t)$ 是 $[0,\ 1;\ 0,\ 1]$ 上的连续函数并且对所有的 $s,\ u \in [0,\ 1]$ 均有 $\int_0^1 K(s, t) K(t, u)\, dt = 0$, 则

$$\int_0^1 K(t, t)\, dt = 0.$$

早先所提的基底问题(参见 Banach)可叙述如下.

我们说 Banach 空间 $\mathbf{X}$ 中的元素序列 $\{x_n\}$ 是基底, 如果每一个元素 $x \in \mathbf{X}$ 可以单值地展成按空间 $\mathbf{X}$ 中的范数收敛的级数

$$x = \sum_{n=1}^{\infty} \lambda_n x_n$$

的话.（基底这个概念是由 J. Schauder [1]引进的. 关于这一概念的详细讨论见 Singer 的论文.）

易见，具有基底的空间一定是可分的. 可分 Hilbert 空间中的完全正交系就是基底(见 IV. 5. 8). 空间 $\mathbf{L}^p(0,1)$、$\mathbf{l}^p(1\leqslant p<\infty)$ 和$\mathbf{C}[0,1]$都是有基底的空间. 例如序列

$$x_n=\{\xi_k\}_{k=1}^{\infty}\qquad\left(\text{其中 }\xi_k=\begin{cases}1, & k=n,\\0, & k\neq n.\end{cases}\right),$$

显然就是空间 $\mathbf{l}^p$ 中的基底.

所谓基底问题就是：是否任何可分 Banach 空间都具有基底？

不难证明：所有具有基底的空间都具有可逼近性质.

上面这两个问题长时期未能得到解决. 只是到了 1972 年瑞典数学家 P. Enflo 构造了一个著名的例子，这个例子立刻对基底问题和可逼近问题作出了否定的回答. 他构造了一个可分自反的 Banach 空间，但是不具有可逼近性，因而也就没有基底. 有关这类问题的详细讨论见 Пелчинский 和 Фигеля 的论文[1].

A. Шанковский 在 1975 年构造了一个可分自反的 Banach 基本空间但不具有可逼近性质的例子.

第十四章　近似方法的一般理论

对于数学分析——微分方程、积分方程、数学物理边值问题、共形映射以及其他问题——的近似解问题已经提出了并且在实际上应用着大量的基于不同思想而设计出来的解法. 例如，对于边值问题，应用变分法和与之类似的方法(Ritz 法，Галеркин 法，矩量法，化成微分方程的方法)，差分法，内插法. 为了讨论应用这些方法的有效性和依据，就必须对它们进行理论研究，在研究时通常提出以下三个问题:

a) 建立算法的可实现性和收敛性;

b) 研究收敛的速度;

c) 误差的有效估计.

这三个问题就其精确程度和困难程度来说后者甚于前者.

对于每一类方程和每一种算法都以自己独有的方法来解决上述问题常常是极其困难的，在很多情况下，直到目前还无法解决.

因此，将这些研究统一起来并建立近似方法的统一理论乃是一个很重要的课题. 解决这个问题的最自然的办法就是利用泛函分析的思想和方法.

在这一章中，我们将要考察对于某一类线性方程建立近似方法的理论. 在这个理论中我们将要考察在各种不同空间(当然是按一定方式联系的)中的精确方程和近似方程，在许多具体方法中，当用有限个的代数方程组代替积分方程时就是一例. 但是，正象我们将要指出的那样，总可以归结成为逼近空间是在其中给出精确方程的那个空间的子空间.

下面有两类定理：一类是根据已知的精确方程可以建立近似方程的可解性及近似解到精确解的收敛性的定理，另一类定理则相反：根据近似解的结果给出精确方程的可解性及精确解与近似解之间接近程度性态的定理.

这一章的结果是由 Л. В. Канторович 在某些其他假定下得到的(见Канторович [9]). 与上面指出的叙述相比 Г. П. Акилов 对 § 1 中某些定理精确化了. 关于利用泛函分析工具建立其他近似方法理论的材料还见于 С. Г. Михлин, М. А. Красносельский, Н. И. Польский, Г. М. Вайникко, М. К. Гавурин 等人的著作(见 Вайникко; Гавурин; Красносельский 等; Красносельский-1; Михлин-II).

§ 1. 关于第二类方程的一般理论

1.1. 设 $\mathbf{X}$ 是赋范空间而 $\tilde{\mathbf{X}}$ 是 $\mathbf{X}$ 的完备子空间(但不假定空间 $\tilde{\mathbf{X}}$ 本身是完备的). 假定存在将空间 $\mathbf{X}$ 投影到 $\tilde{\mathbf{X}}$ 上的连续线性算子 P, 也就是说

$$P(\mathbf{X})=\tilde{\mathbf{X}},\quad P^2=P.$$

易见, 算子 P 不改变 $\tilde{\mathbf{X}}$ 中的元素.

例如, 设 $\mathbf{X}=\mathbf{C}[a,b]$ 而 $\tilde{\mathbf{X}}$ 是所有次数不高于 $(n-1)$ 的多项式的全体. 算子 P 把每个连续函数 $x\in\mathbf{X}=\mathbf{C}[a,b]$ 和它的按预先给定的结点组 $t_1, t_2, \cdots, t_n$ 构成的插值多项式对应起来.

其次, 我们再考察两个方程, 第一个是空间 $\mathbf{X}$ 中的方程

$$Kx\equiv x-\lambda Hx=y \tag{1}$$

第二个是空间 $\tilde{\mathbf{X}}$ 中的方程

$$\tilde{K}\tilde{x}\equiv\tilde{x}-\lambda\tilde{H}\tilde{x}=Py. \tag{2}$$

上面的 H 是 $\mathbf{X}$ 中的连续线性算子而 $\tilde{H}$ 是 $\tilde{\mathbf{X}}$ 中的连续线性算子. 方程(1)叫精确方程而方程(2)叫相应于精确方程(1)的近似方程.

在以后的叙述中, 我们假定空间 $\mathbf{X}$ 和 $\tilde{\mathbf{X}}$, 算子 H 和 $\tilde{H}$ 由下列

条件建立联系:

I(算子H和$\widetilde{H}$的可接近性条件). 对于任意的$\tilde{x}\in\widetilde{\mathbf{X}}$

$$\|PH\tilde{x}-\widetilde{H}\tilde{x}\|\leqslant\eta\|\tilde{x}\|.$$

这个关系式等价于

$$\|PK\tilde{x}-\widetilde{K}\tilde{x}\|\leqslant|\lambda|\eta\|\tilde{x}\| \qquad (\tilde{x}\in\widetilde{\mathbf{X}}).$$

II(形如Hx的元素可由$\widetilde{\mathbf{X}}$中的元素很好地逼近的条件). 对任何$x\in\mathbf{X}$, 可找到$\tilde{x}\in\widetilde{\mathbf{X}}$使得

$$\|Hx-\tilde{x}\|\leqslant\eta_1\|x\|.$$

III(精确方程的自由项可很好得以逼近的条件). 存在元素$\tilde{y}\in\widetilde{\mathbf{X}}$使得

$$\|y-\tilde{y}\|\leqslant\eta_2\|y\|.$$

和前面的两个条件不同, 这里的η_2一般地依赖于y.

1.2. 对于满足上述条件(或其中部分条件)的方程, 我们要证明一系列建立了精确方程和近似方程之间联系的定理.

定理 1 (关于近似方程的可解性). 设条件 I 和 II 成立而且算子K具有双向的连续逆算子. 这时, 如果

$$q=|\lambda|[\eta+\|I-P\|\eta_1]\|K^{-1}\|<1, \tag{3}$$

则算子$\widetilde{K}$也有连续逆算子$\widetilde{K}^{-1}$, 并且

$$\|\widetilde{K}^{-1}\|\leqslant\frac{\|K^{-1}\|}{1-q}. \tag{4}$$

证. 首先假定空间$\mathbf{X}$是完备空间. 考察算子$K_1=I-\lambda PH$. 对于任意的元素$x\in\mathbf{X}$, 由条件 II 作$\tilde{x}\in\widetilde{\mathbf{X}}$使得

$$\|Hx-\tilde{x}\|\leqslant\eta_1\|x\|.$$

这时

$$\begin{aligned}\|(K-K_1)x\|&=|\lambda|\,\|Hx-PHx\|\\&=|\lambda|\,\|Hx-\tilde{x}+P\tilde{x}-PH\tilde{x}\|\\&=|\lambda|\,\|(I-P)(Hx-\tilde{x})\|\leqslant|\lambda|\,\|I-P\|\eta_1\|x\|.\end{aligned}$$

注意元素 x 的任意性，则我们证得了

$$\|K-K_1\|\leqslant|\lambda|\,\|I-P\|\eta_1.$$

算子 K_1 可以表示成

$$K_1=K(I-K^{-1}(K-K_1)),\tag{5}$$

并且算子 $K^{-1}(K-K_1)$ 满足如下估计

$$\|K^{-1}(K-K_1)\|\leqslant|\lambda|\,\|I-P\|\eta_1\|K^{-1}\|\leqslant q<1.$$

按 Banach 定理（定理 V. 4. 3）存在逆算子 $(I-K^{-1}(K-K_1))^{-1}$，并且

$$\|(I-K^{-1}(K-K_1))^{-1}\|\leqslant\frac{1}{1-|\lambda|\,\|I-P\|\eta_1\|K^{-1}\|}.$$

从表达式(5)可推知算子K_1可逆，$K_1^{-1}=(I-K^{-1}(K-K_1))^{-1}K^{-1}$，所以

$$\|K_1^{-1}\|\leqslant\frac{\|K^{-1}\|}{1-|\lambda|\,\|I-P\|\eta_1\|K^{-1}\|}.\tag{6}$$

在空间$\tilde{\mathbf{X}}$中考察算子 $\tilde{K}_1\tilde{x}=\tilde{x}-\lambda PH\tilde{x}$. 易见，对于任意的$\tilde{x}\in\tilde{\mathbf{X}}$ 有 $\tilde{K}_1\tilde{x}=K_1\tilde{x}$. 算子 K_1 还具有性质：如果 $\tilde{y}\in\tilde{\mathbf{X}}$ 则 $K_1^{-1}\tilde{y}\in\tilde{\mathbf{X}}$. 事实上，如果 $x'=K_1^{-1}\tilde{y}$，则 $x'-PHx'=\tilde{y}$，$x'=\tilde{y}+PHx'\in\tilde{\mathbf{X}}$. 因此，算子$\tilde{K}_1$具有连续逆，并且在$\tilde{\mathbf{X}}$上与 K_1^{-1} 重合，且

$$\|\tilde{K}_1^{-1}\|\leqslant\|K_1^{-1}\|.\tag{7}$$

我们现在来估计差 $\tilde{K}-\tilde{K}_1$. 根据条件 I，对于任意的 $\tilde{x}\in\tilde{\mathbf{X}}$，有

$$\|\tilde{K}\tilde{x}-\tilde{K}_1\tilde{x}\|=|\lambda|\,\|PH\tilde{x}-\tilde{H}\tilde{x}\|\leqslant|\lambda|\eta\|\tilde{x}\|,\tag{8}$$

所以

$$\|\tilde{K}-\tilde{K}_1\|\leqslant|\lambda|\eta.\tag{9}$$

因此，根据(6)，(7)和(9)

$$\begin{aligned}\|\tilde{K}_1^{-1}\|\,\|\tilde{K}-\tilde{K}_1\|&\leqslant\frac{\|K^{-1}\|\,|\lambda|\eta}{1-|\lambda|\,\|I-P\|\eta_1\|K^{-1}\|}\\&=1-\frac{1-q}{1-|\lambda|\,\|I-P\|\eta_1\|K^{-1}\|}<1.\end{aligned}\tag{10}$$

把 $\widetilde{K}$ 写成 $\widetilde{K}=\widetilde{K}_1(I-\widetilde{K}_1^{-1}(\widetilde{K}_1-K))$，再应用 Banach 定理，我们就证得了逆算子 $\widetilde{K}^{-1}$ 的存在性，且有估计式

$$\|\widetilde{K}^{-1}\|\leqslant\frac{\|\widetilde{K}_1^{-1}\|}{1-\|\widetilde{K}_1^{-1}\|\|\widetilde{K}-\widetilde{K}_1\|}$$

$$\leqslant\frac{(1-|\lambda|\|I-P\|\eta_1\|K^{-1}\|)\|K_1^{-1}\|}{1-q}$$

$$\leqslant\frac{\|K^{-1}\|}{1-q}. \tag{11}$$

至此对完备空间 $\mathbf{X}$ 定理已得证.

当空间 $\mathbf{X}$ 不完备时，容易把它归结为已考察过的情形. 事实上，如果引进空间 $\mathbf{X}$ 的完备化空间，则算子 K 和 K^{-1} 的闭包(参见定理 V.8.2)是互逆的，而算子 H 的闭包与 $\widetilde{H}$ 的闭包之间亦满足条件 I 和 II. 这时，条件中的常数 η 是原来的而 η_1 则是和原来的那个常数充分接近的一个新的常数.

1.3. 近似方程(2)的解自然看成是精确方程(1)的近似解. 此近似解的误差估计由下述定理给出.

定理 2 (关于近似解的误差估计). 如果条件 I、II 和 III 成立，存在连续算子 $\widetilde{K}^{-1}$(特别，如果定理 1 的条件成立)及方程(1)有解 x^*，则估计式

$$\|x^*-\tilde{x}^*\|\leqslant p\|x^*\| \tag{12}$$

成立，这里的 $\tilde{x}^*$ 是方程(2)的解，而

$$p=2|\lambda|\eta\|\widetilde{K}^{-1}\|+(\eta_1|\lambda|+\eta_2\|K\|)(1+\|\widetilde{K}^{-1}PK\|). \tag{13}$$

证. 首先证明定理的条件使得可用 $\widetilde{\mathbf{X}}$ 中的元素逼近 x^* 到 $(\eta_1+\eta_2)$ 阶的程度. 详而言之，即证明存在元素 $\tilde{x}\in\widetilde{\mathbf{X}}$ 使得

$$\|x^*-\tilde{x}\|\leqslant\varepsilon\|x^*\|, \tag{14}$$

其中

$$\varepsilon=\min[1,\eta_1|\lambda|+\eta_2|K|]. \tag{15}$$

事实上，由条件 II 和 III，可取 $\tilde{y}\in\widetilde{\mathbf{X}}$ 和 $\tilde{z}\in\widetilde{\mathbf{X}}$ 使得

$$\|Hx^*-\tilde{z}\|\leqslant\eta_1\|x^*\|,\ \|y-\tilde{y}\|\leqslant\eta_2\|y\|\leqslant\eta_2\|K\|\|x^*\|,$$

再令 $\tilde{x}=\lambda\tilde{z}+\tilde{y}$, 则

$$\|x^*-\tilde{x}\|=\|\lambda Hx^*+y-(\lambda\tilde{z}+\tilde{y})\|\leqslant(\eta_1|\lambda|+\eta_2\|K\|)\|x^*\|.$$

另一方面, 若命 $\tilde{x}=\mathbf{0}$, 则当 $\varepsilon=1$ 时(14)式成立. 由此推知, 可取(15)式作为 ε.

现在转向不等式(12)的证明.

记 $\tilde{x}_0=\widetilde{K}^{-1}PK\tilde{x}$, 则

$$\|x^*-\tilde{x}^*\|\leqslant\|x^*-\tilde{x}\|+\|\tilde{x}-\tilde{x}_0\|+\|\tilde{x}_0-\tilde{x}^*\|. \tag{16}$$

我们分别来估计每一项. 第一项的估计已经有了, 这就是(14)式. 第二项可这样来估计: 由条件 I, 有

$$\begin{aligned}\|\tilde{x}-\tilde{x}_0\|&=\|\widetilde{K}^{-1}\widetilde{K}\tilde{x}-\widetilde{K}^{-1}PK\tilde{x}\|\\&\leqslant\|K^{-1}\|\|\widetilde{K}\tilde{x}-PK\tilde{x}\|\\&\leqslant|\lambda|\eta\|\widetilde{K}^{-1}\|\|\tilde{x}\|.\end{aligned} \tag{17}$$

但由(14)式

$$\|\tilde{x}\|\leqslant\|x^*\|+\|x^*-\tilde{x}\|\leqslant(1+\varepsilon)\|x^*\|.$$

将这个式子代入(17), 最后得到

$$\|\tilde{x}-\tilde{x}_0\|\leqslant|\lambda|\eta(1+\varepsilon)\|\widetilde{K}^{-1}\|\|x^*\|. \tag{18}$$

注意到 x^* 是方程(2)的解, 因而

$$\tilde{x}^*=\widetilde{K}^{-1}PKx^*.$$

我们来估计(16)式中最后一项. 由(14)可知

$$\begin{aligned}\|\tilde{x}_0-\tilde{x}^*\|&=\|\widetilde{K}^{-1}PK\tilde{x}-\widetilde{K}^{-1}PKx^*\|\\&\leqslant\|\widetilde{K}^{-1}PK\|\|\tilde{x}-x^*\|\leqslant\varepsilon\|\widetilde{K}^{-1}PK\|\|x^*\|.\end{aligned}$$

此式与(14), (18)一起就给出如下的估计式

$$\|x^*-\tilde{x}^*\|\leqslant[|\lambda|\eta(1+\varepsilon)\|\widetilde{K}^{-1}\|+\varepsilon(1+\|\widetilde{K}^{-1}PK\|)]\|x^*\|. \tag{19}$$

只须取 $1+\varepsilon\leqslant2$, $\varepsilon\leqslant\eta_1|\lambda|+\eta_2\|K\|$, 即知(12)成立.

定理证毕.

注 1. 可以证明，用$\tilde{\mathbf{X}}$中的元素去逼近元素 x^* 的可能性不必根据 II 和 III 而可以直接建立起来. 这时仍是先用到不等式(19)，自然，此时不必假定条件 II 和 III 满足.

指出这一点是有用的，因为这里讲的是逼近具体元素 x^* 而且 ε 可以依赖于它，而在条件 II 中 η_1 不依赖于元素 x.

注 2. 由定理 2 也容易得到近似解接近精确解的估计而不一定要用精确解来进行估计. 也就是说，如果在定理 2 的条件中有 $p<1$，则成立如下的估计式

$$\|x^*-\tilde{x}^*\|\leqslant\frac{p}{1-p}\|\tilde{x}^*\|. \tag{20}$$

事实上，

$$\|x^*\|\leqslant\|\tilde{x}^*\|+\|x^*-\tilde{x}^*\|.$$

将它代入不等式(12)得到

$$\|x^*-\tilde{x}^*\|\leqslant p\|\tilde{x}^*\|+p\|x^*-\tilde{x}^*\|,$$

从这个不等式中解出$\|x^*-\tilde{x}^*\|$即得(20)式.

1. 4. 现在设有一个近似方程序列及由它所得到的近似解. 在这种情形，空间$\tilde{\mathbf{X}}$、算子 $\tilde{\Pi}(\tilde{K})$, P 以及常数 $\eta, \eta_1, \eta_2, q, p, \varepsilon, \cdots$ 实际上都依赖于附标 n，但为了简单起见，我们将不记其附标.

下述定理给出近似解序列$\{\tilde{x}_n^*\}$收敛于精确解的条件.

定理 3. 如果下列条件成立:

1) 算子K具有连续的逆算子;

2) 对每个 $n=1, 2, \cdots$, 条件 I, II, III 成立，并且

$$\lim_{n\to\infty}\eta=0,\quad \lim_{n\to\infty}\eta_1\|P\|=0,\quad \lim_{n\to\infty}\eta_2\|P\|=0, \tag{21}$$

则对于充分大的 n 近似方程可解并且近似解序列收敛于精确解:

$$\lim_{n\to\infty}\|x^*-\tilde{x}_n^*\|=0.$$

确切地说，有

$$\|x^*-\tilde{x}_n^*\|\leqslant Q_0\eta+Q_1\eta_1\|P\|+Q_2\eta_2\|P\|, \tag{22}$$

其中 Q_0, Q_1 和 Q_2 是某些常数.

证. 因为$\|P\|\geqslant 1$, 故由(21)推知 $\lim\limits_{n\to\infty}\eta_1=0$, 因此当 n 充分大时在定理 1 中将会有 $q<1/2$. 这样一来, 对于此 n 存在连续算子 $\widetilde{K}^{-1}$ 且

$$\|\widetilde{K}^{-1}\|\leqslant\frac{\|K^{-1}\|}{1-q}<2\|K^{-1}\|.$$

由此可见$\|\widetilde{K}^{-1}\|$有界并且与 n 无关. 据此, 根据定理 2 中的估计式(12)即得到要证的(22).

注. 如果在证明中用估计式(19)来代替(12), 则得

$$\|x^*-\tilde{x}_n^*\|\leqslant Q'\eta+Q''\varepsilon\|P\| \tag{23}$$

且收敛条件(21)就变成

$$\lim_{n\to\infty}\eta=0,\quad \lim_{n\to\infty}\varepsilon\|P\|=0. \tag{24}$$

这时条件 III 是多余的(至于条件 II, 它是必要的, 因为在定理 1 中要用到它. 并且, 条件 $\eta_1\|P\|\to 0$ 可用要求当 n 充分大时 $q\leqslant q_0<1$ 来代替).

1.5. 定理 1 给出根据精确方程的可解性判断近似方程可解性的可能性. 下面的定理给出了相反的论断.

定理 4. 如果算子 $\widetilde{K}$ 具有满足条件 I 和 II 的连续逆算子, 并且

$$r=|\lambda|\eta(1+|\lambda|\eta_1)\|\widetilde{K}^{-1}\|+|\lambda|\eta_1(1+\|\widetilde{K}^{-1}PK\|)<1, \tag{25}$$

则算子 K 具有连续的左逆算子, 这个逆算子的范数有如下的估计

$$\|K^{-1}\|\leqslant\frac{1+\|\widetilde{K}^{-1}P\|+|\lambda|\eta\|\widetilde{K}^{-1}\|+\|\widetilde{K}^{-1}PK\|}{1-r}. \tag{26}$$

证. 设 x^* 是空间 $\mathbf{X}$ 中的任意非零元素. 它显然是方程

$$Kx=Kx^*$$

的解.

根据条件 II, 可找到元素 $\tilde{x}\in\widetilde{\mathbf{X}}$ 使得

$$\|Hx^* - \tilde{x}\| \leqslant \eta_1 \|x^*\|.$$

因为 $x^* = \lambda Hx^* + Kx^*$, 所以

$$\|x^* - \lambda\tilde{x}\| \leqslant |\lambda|\eta_1\|x^*\| + \|Kx^*\| = \left(|\lambda|\eta_1 + \frac{\|Kx^*\|}{\|x^*\|}\right)\|x^*\|.$$

这个不等式使得我们可以利用具有如下 ε 值的定理 2(定理 2 的注 1 的形式):

$$\varepsilon = |\lambda|\eta_1 + \frac{\|Kx^*\|}{\|x^*\|}. \tag{27}$$

因为对于我们所考察的精确方程的近似方程是

$$\widetilde{K}\tilde{x} = PKx^*,$$

它的解是 $\tilde{x}^* = \widetilde{K}^{-1}PKx^*$, 所以将(27)式的 ε 值代入(19)式得到

$$\|x^* - \widetilde{K}^{-1}PKx^*\| \leqslant \Big[|\lambda|\eta\left(1 + |\lambda|\eta_1 + \frac{\|Kx^*\|}{\|x^*\|}\right)\|\widetilde{K}^{-1}\|$$
$$+ \left(|\lambda|\eta_1 + \frac{\|Kx^*\|}{\|x^*\|}\right)(1 + \|\widetilde{K}^{-1}PK\|)\Big]\|x^*\|$$
$$= [|\lambda|\eta(1 + |\lambda|\eta_1)\|\widetilde{K}^{-1}\| + |\lambda|\eta_1(1 + \|\widetilde{K}^{-1}PK\|)]\|x^*\|$$
$$+ [|\lambda|\eta\|\widetilde{K}^{-1}\| + 1 + \|\widetilde{K}^{-1}PK\|]\|Kx^*\|.$$

由此

$$\|x^*\| \leqslant \|x^* - \widetilde{K}^{-1}PKx^*\| + \|\widetilde{K}^{-1}P\|\|Kx^*\|$$
$$\leqslant r\|x^*\| + [1 + \|\widetilde{K}^{-1}P\| + |\lambda|\eta\|\widetilde{K}^{-1}\|$$
$$+ \|\widetilde{K}^{-1}PK\|]\|Kx^*\|$$

或

$$\|Kx^*\| \geqslant \frac{1 - r}{1 + \|\widetilde{K}^{-1}P\| + |\lambda|\eta\|\widetilde{K}^{-1}\| + \|\widetilde{K}^{-1}PK\|}\|x^*\| = Q\|x^*\|.$$

因为 $r < 1$, 故 $Q > 0$, 再由 V.4.4 中所述的结果即证得了该定理.

推论 如果算子 K 满足条件:

(A) 从方程(1)的解的唯一性可推出对任意的右端项的可解性.

这时在定理的条件下，算子 K 存在双向连续逆算子 K^{-1}.

1.6. 如果条件 III 仅对充分大的常数 η_2 满足，也就是说如果方程(1)的右端不能由 $\widetilde{X}$ 中的元素很好地逼近，则也不能很好地逼近解 x^* 而近似解 $\tilde{x}^*$ 将明显变坏. 这时将自由项加以"正则化"是适当的. 对方程 (1) 作自变量的代换 $x=y+z$. 关于 z 的新方程为

$$Kz=y' \qquad (y'=\lambda Hy). \tag{28}$$

由条件 II, 这个方程的右端 y' 已经是"正则"的了.

关于方程(28)的近似方程是

$$\widetilde{K}\tilde{z}=\lambda PHy. \tag{29}$$

以 $\tilde{z}^*$ 表示此方程的解, 自然将元素

$$x'=y+\tilde{z}^* \tag{30}$$

作为方程(1)的近似解. 因为

$$\|x^*-x'\|=\|z^*-\tilde{z}^*\|,$$

其中 z^* 表示方程(28)的解, 则近似解(30)的误差估计式可由定理 2 得到. 不过这时作为 η_2 应取为 $|\lambda|\eta_1$*). 但直接考察所给的情形, 我们还可得到更精确的估计.

定理 5. 如果条件 I 和 II 成立并且 $\widetilde{K}$ 存在连续的逆算子 $\widetilde{K}^{-1}$ 和方程(1)有解 x^*, 则有如下的估计式

$$\|x^*-x'\|\leqslant p'\|x^*\| \tag{31}$$

其中

$$p'=|\lambda|[\eta_1(1+\|\widetilde{K}^{-1}PK\|)+|\lambda|\eta\|\widetilde{K}^{-1}\|(\|H\|+\eta_1)] \tag{32}$$

证. 因为

$$z^*=\lambda Hz^*+\lambda Hy=\lambda Hx^*, \tag{33}$$

据条件 II, 可求得元素 $\tilde{z}\in\widetilde{X}$ 使得

$$\|z^*-\tilde{z}\|\leqslant|\lambda|\eta_1\|x^*\|. \tag{34}$$

记

$$\tilde{z}_0=\widetilde{K}^{-1}PK\tilde{z}.$$

---

*) 或者更小的量 $|\lambda|\eta_1^*$, 只要对某个 $\tilde{y}\in\widetilde{X}$ 不等式

$$\|Hy-\tilde{y}\|\leqslant\eta_1^*\|y\|$$

成立即可.

因此, 如果条件(21)中的第二个成立, 则第三个条件总能得到满足.

因为

$$\widetilde{K}\tilde{z}^* = PKx^*,$$

则

$$\tilde{z}^* = \widetilde{K}^{-1}PKz^*,$$

由此得到

$$\|\tilde{z}^* - \tilde{z}_0\| \leqslant \|\widetilde{K}^{-1}PK\|\|z^* - \tilde{z}\| \leqslant |\lambda|\eta_1\|\widetilde{K}^{-1}PK\|\|x^*\|. \tag{35}$$

其次，由条件 I, (33) 和 (34)

$$\begin{aligned}\|\tilde{z}^* - \tilde{z}_0\| &\leqslant \|\widetilde{K}^{-1}\|\|\widetilde{K}\tilde{z} - PK\tilde{z}\| \leqslant |\lambda|\eta\|\widetilde{K}^{-1}\|\|\tilde{z}\| \\ &\leqslant |\lambda|\eta\|\widetilde{K}^{-1}\|[\|z^*\| + \|\tilde{z} - z^*\|] \\ &\leqslant |\lambda|\eta\|\widetilde{K}^{-1}\|(|\lambda|\|H\| + |\lambda|\eta_1)\|x^*\|.\end{aligned}$$

但由 (34) 和 (35) 得

$$\begin{aligned}\|x^* - x'\| &\leqslant \|z^* - \tilde{z}\| + \|\tilde{z} - \tilde{z}_0\| + \|\tilde{z}_0 - \tilde{z}^*\| \\ &\leqslant [|\lambda|\eta_1 + |\lambda|^2\eta\|\widetilde{K}^{-1}\|(\|H\| + \eta_1) + |\lambda|\eta_1\|\widetilde{K}^{-1}PK\|]\|x^*\| \\ &= p'\|x^*\|,\end{aligned}$$

这正是所要证明的.

在这种情况下，用类似的办法可使得定理 4 的结果精确化.

1.7. 我们现在来考察算子 $\widetilde{H}$ 的特征值收敛于算子 H 的特征值的问题. 我们只限于对 H 和 $\widetilde{H}$ 均为紧算子的情形进行讨论，因而其特征集是离散的.

首先，在定理 3 的假定之下可以证明算子 $\widetilde{H}$ 的特征值只能收敛于算子 H 的特征值. 事实上，在 λ 平面上取以坐标原点为中心以 R 为半径的圆 C_0，并设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是算子 H 的落在 C_0 中的特征值. 用 D 表示从 C_0 中除去点 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 的 δ 邻域后所得到的闭域. 在此闭域中算子 K_λ^{-1} ($K_\lambda = I - \lambda H$) 存在且 $\|K_\lambda^{-1}\|$ 作为 λ 的函数是有界的. 这时，由定理 3 可知对充分大的 n, $\|\widetilde{K}_\lambda^{-1}\|$ ($\widetilde{K}_\lambda = \widetilde{I} - \lambda\widetilde{H}$) 也有界且不依赖于 n. 因而在上述条件下算子 $\widetilde{H}$ 的特征值只可能在除掉的 δ 邻域中或在 C_0 外找到. 由此显然可见若当 $n \to \infty$ 时这些特征值有有限的极限，则这个极限必是算子 H 的一个特征值.

同时还可证明，算子 H 的每一个特征值都是算子 $\widetilde{H}$ 的特征值的极限. 事实上，如果某个 λ_k 不是这种极限，则在点 λ_k 的以圆周 C_k 为边界的某个邻域中当 n 充分大时没有算子 $\widetilde{H}$ 的特征值. 函数 $\|K_\lambda^{-1}\|$ 在圆周 C_k 上有界. 由 (11)，在此圆周上 $\|\widetilde{K}_\lambda^{-1}\|$ 亦有界. 这时，在此圆周内 $\|\widetilde{K}_\lambda^{-1}\|$ 以同一个数作为界. 应用解析函数的极大模原理于函数 $\psi(\lambda) = f(\widetilde{K}_\lambda^{-1}\tilde{x})$ (这里的 f 是 $\mathbf{X}$ 中的任

意泛函而 $\tilde{x}$ 是 $\tilde{\mathbf{X}}$ 中任意元素）可知 $\|\tilde{K}_\lambda^{-1}\|$ 在 C_k 的内部更是有界的了. 利用定理 4, 我们得知 K_λ^{-1} 在所有邻域中都存在, 从而导致矛盾.

在近似方法的一般理论中, Троичкая[1]和 Вайникко 曾考察过特征值和特征元的收敛问题. 这一问题还可参见 Михлин-II.

1.8. 我们指出一种重要的情形, 即可将以前所证明的定理加以简化的情形.

近似方程(2)通常是用特殊的方法建立的, 即作为算子 $\tilde{H}$ 是取作 PH 的(算子 H 在这里应认为是从空间 $\tilde{\mathbf{X}}$ 到 $\mathbf{X}$ 的算子); 因此, 近似方程就是

$$\tilde{x}-\lambda PH\tilde{x}=Py \qquad (\tilde{K}=PK). \tag{36}$$

在近似算子 $\tilde{H}$ 这样的选择之下, 条件 I 显然在 $\eta=0$ 时成立. 这可使定理的形式化简. 例如在定理 1 中应取

$$q=|\lambda|\eta_1\|I-P\|\|K^{-1}\|. \tag{37}$$

在定理 2 中应取

$$p=(\eta_1|\lambda|+\eta_2\|K\|)(1+\|\tilde{K}^{-1}PK\|). \tag{38}$$

定理 3 中的条件(21)应代之以

$$\lim_{n\to\infty}\eta_1\|P\|=\lim_{n\to\infty}\eta_2\|P\|=0. \tag{39}$$

在定理 4 中

$$r=|\lambda|\eta_1(1+\|K^{-1}PK\|), \qquad \|K^{-1}\|\leqslant\frac{1+\|\tilde{K}^{-1}P\|+\|\tilde{K}^{-1}PK\|}{1-r}. \tag{40}$$

我们现在来给出一些条件, 这些条件可以保证关系式(39)成立从而可使近似解序列收敛于精确解. 设 P_n 是从 $\mathbf{X}$ 到子空间 $\tilde{\mathbf{X}}_n$ 上的投影算子序列, 每个 P_n 对应于它自己的近似方程(36).

定理 6. 如果下列条件被满足:

1) $\mathbf{X}$ 是完备空间;

2) 在 $\mathbf{X}$ 上 $P_n\to I$, 即 $\lim\limits_{n\to\infty}P_nx=x \qquad (x\in\mathbf{X})$;

3) H是紧算子,

则在条件 II 和 III 中所指定的常数$\eta_1^{(n)}$和$\eta_2^{(n)}$可以这样来选择:

$$\eta_1^{(n)}\to 0,\quad \eta_2^{(n)}\to 0 \qquad (n\to\infty).$$

证. 设B是空间$\mathbf{X}$中的单位球. 集合$H(B)$是紧的, 因而由 Гель фанд 定理(IX.1.4)知在这个集合上算子序列P_n一致地收敛于单位算子I. 命

$$\eta_1^{(n)}=\sup_{z\in H(B)}\|P_n z-z\| \qquad (n=1, 2, \cdots),$$

则

$$\eta_1^{(n)}\to 0.$$

另一方面, 对任何$x\in\mathbf{X}$, 均有

$$\|P_nHx-Hx\|\leqslant\eta_1^{(n)}\|x\|.$$

由于$P_nHx\in\tilde{\mathbf{X}}_n$, 故由这个不等式推知$\eta_1^{(n)}$满足条件 II.

作为$\eta_2^{(n)}$可取为

$$\eta_2^{(n)}=\frac{\|P_ny-y\|}{\|y\|},$$

它显然趋于零.

推论 1. 在定理的条件下, 如果λ不是算子H的特征值, 则定理 3 的条件成立, 因而近似解序列$\{\tilde{x}_n^*\}$收敛于精确解.

事实上, 由定理的条件 2) 可知$\sup\limits_n\|P_n\|<\infty$(见 VII.1.2), 故(39)式成立.

推论 2. 算子H的特征值是算子$\tilde{H}_n$的特征值序列的极限.

我们现在来更详细讨论当$\tilde{\mathbf{X}}$是有限维(n 维)空间的情形. 这时, 每一个$\tilde{x}\in\tilde{\mathbf{X}}$可唯一地表示成

$$\tilde{x}=c_1\omega_1+c_2\omega_2+\cdots+c_n\omega_n, \tag{41}$$

其中元素$\omega_1, \omega_2, \cdots, \omega_n$构成$\tilde{\mathbf{X}}$的基.

设$f_1, f_2, \cdots, f_n$是$\tilde{\mathbf{X}}$中线性泛函的完全组, 即由等式

$$f_j(\tilde{x})=0 \qquad (j=1,2,\cdots,n)$$

可推出 $\tilde{x}=\mathbf{0}$ 的泛函组.

近似方程(36)等价于

$$f_j(PK\tilde{x})=f_j(Py) \qquad (j=1,2,\cdots,n).$$

求方程(36)的形如(41)的解,就得到关于系数 c_k 的线性代数方程组

$$\sum_{k=1}^{n}\alpha_{jk}c_k-\lambda\sum_{k=1}^{n}a_{jk}c_k=b_j \qquad (j=1,2,\cdots,n), \tag{42}$$

其中

$$\alpha_{jk}=f_j(\omega_k), \quad a_{jk}=f_j(PH\omega_k), \quad b_j=f_j(Py)$$
$$(j,k=1,2,\cdots,n).$$

如果泛函组 $f_1,f_2,\cdots,f_n$ 与基 $\omega_1,\omega_2,\cdots,\omega_n$ 双正交,则(42)可简化为

$$c_j-\lambda\sum_{k=1}^{n}a_{jk}c_k=b_j \qquad (j=1,2,\cdots,n). \tag{43}$$

特别,如果 $\mathbf{X}$ 是 Hilbert 空间,P 是正交投影算子,设 $\omega_1,\omega_2,\cdots,\omega_n$ 是正交组,则得到

$$\alpha_{jk}=f_j(PH\omega_k)=(PH\omega_k,\omega_j)=(H\omega_k,P\omega_j)=(H\omega_k,\omega_j)$$
$$(j,k=1,2,\cdots,n).$$
$$b_j=f_j(Py)=(Py,\omega_j)=(y,P\omega_j)=(y,\omega_j)$$
$$(j=1,2,\cdots,n),$$

于是,方程(42)可写成

$$c_j-\lambda\sum_{k=1}^{n}c_k(H\omega_k,\omega_j)=(y,\omega_j) \qquad (j=1,2,\cdots,n). \tag{44}$$

方程组(42)叫做抽象形式的 Галеркин 法*).

*) 对于非线性方程的 Галеркин 法,在 Красносельский 等和 Вайникко 的书中作过讨论.

我们再指出一般理论的另一种特殊情形，这时还可以作很大的简化.

如果空间$\mathbf{X}$与$\widetilde{\mathbf{X}}$重合，则条件II和III显然成立，且$\eta_1=\eta_2=0$. 这时条件I变为算子H和$\widetilde{H}$接近这一条件(此时，这两个算子作用于同一个空间)，即条件

$$\|H-\widetilde{H}\|\leqslant\eta.$$

但应当指出，这种情形从一般理论的观点来看并不是有很大兴趣的，因为这时所有理论均可用初等工具加以证明(见V. 4.6).

1.9. 经常有这样的情形，即某种方法的近似方程不是在空间$\mathbf{X}$的子空间中讨论而是在另一个通常与基本空间$\mathbf{X}$的某个子空间同构的空间$\overline{\mathbf{X}}$上来讨论.

我们认为完备子空间$\widetilde{\mathbf{X}}\subset\mathbf{X}$由将$\widetilde{\mathbf{X}}$单值地映射到完备空间$\overline{\mathbf{X}}$上的连续线性算子$\varphi_0$来确定. 这就保证了逆算子$\varphi_0^{-1}$的连续性. 假设存在连续线性算子$\varphi$($\varphi$是$\varphi_0$在整个$\mathbf{X}$上的扩张)，因此，$\varphi$是从$\mathbf{X}$到$\overline{\mathbf{X}}$上的连续线性算子，它在$\widetilde{\mathbf{X}}$上与$\varphi_0$重合.

若将P看成是如在1.1中所引进的从$\mathbf{X}$到$\widetilde{\mathbf{X}}$的投影算子，则作为φ可取为

$$\varphi=\varphi_0P. \tag{45}$$

这个等式两端乘以φ_0^{-1}后得到

$$P=\varphi_0^{-1}\varphi. \tag{46}$$

由于空间$\widetilde{\mathbf{X}}$和$\overline{\mathbf{X}}$元素之间存在一一对应，故可把近似方程(2)变换成在空间$\overline{\mathbf{X}}$中的等价方程. 为此，可在(2)中作代换$\tilde{x}=\varphi_0^{-1}\bar{x}$，然后再在方程的两端作用算子$\varphi_0$. 用这种办法变换得到的方程是

$$\bar{x}-\varphi_0\widetilde{H}\varphi_0^{-1}\bar{x}=\varphi_0Py.$$

用$\overline{H}$表示从$\overline{\mathbf{X}}$到自身的算子:

$$\overline{H}=\varphi_0\widetilde{H}\varphi_0^{-1}, \tag{47}$$

再注意到(45)，则我们可将近似方程改写为

$$\overline{K}\bar{x} \equiv \bar{x} - \lambda \overline{H}\bar{x} = \varphi y \qquad (\overline{K} = \varphi_0 \widetilde{K} \varphi_0^{-1}). \tag{48}$$

由于在应用中常会遇到近似方程是用(48)式给出的情形，所以我们总是用算子 $\overline{H}$，φ 和 $\overline{K}$ 的语言来给出一般理论中基本定理的叙述．为此，应该用(46)式来代替式中的 P，对于算子 $\widetilde{H}(\widetilde{K})$，从(47)可得

$$\widetilde{H} = \varphi_0^{-1}\overline{H}\varphi_0 \qquad (\widetilde{K} = \varphi_0^{-1}\overline{K}\varphi_0). \tag{49}$$

我们来建立条件 I．使用新的表示法，它可写成

$$\|\varphi_0^{-1}\overline{H}\varphi_0\tilde{x} - \varphi_0^{-1}\varphi H\tilde{x}\| \leqslant \eta\|\tilde{x}\| \qquad (\tilde{x} \in \widetilde{\mathbf{X}}).$$

这个不等式在下述条件下成立：

Ia．对于任何 $\tilde{x} \in \widetilde{\mathbf{X}}$，成立不等式

$$\|\overline{H}\varphi_0\tilde{x} - \varphi H\tilde{x}\| \leqslant \overline{\eta}\|\tilde{x}\|.$$

这时作为条件 I 中的 η 应取

$$\eta = \overline{\eta}\|\varphi_0^{-1}\|. \tag{50}$$

条件 II 和 III 不涉及到算子 $\widetilde{H}$ 和 P，因此在新的情形下它们不改变．

现在给出定理 1 的新的形式．

定理 1a．设条件 Ia 和 II 满足，并且算子 K 具有连续逆算子．这时，如果

$$\bar{q} = |\lambda|[\overline{\eta}\|\varphi_0^{-1}\| + \|I - \varphi_0^{-1}\varphi\|\eta_1]\|K^{-1}\| < 1 \tag{51}$$

则算子 $\overline{K}$ 也具有连续逆算子 $\overline{K}^{-1}$，并且

$$\|\overline{K}^{-1}\| \leqslant \frac{\overline{N}}{1-\bar{q}}, \tag{52}$$

其中

$$\overline{N} = \|K^{-1}\|\|\varphi_0\|\|\varphi_0^{-1}\|. \tag{53}$$

事实上，$\overline{K} = \varphi_0\widetilde{K}\varphi_0^{-1}$，由定理 1 可知对算子 $\widetilde{K}$ 存在逆算子 $\widetilde{K}^{-1}$，且对此逆算子估计式(4)成立．

如果用 $\bar{x}^*$ 表示方程(48)的解，则方程(1)的近似解 $\tilde{x}^*$将为 $\tilde{x}^*=\varphi_0^{-1}\bar{x}^*$. 近似解的误差估计可经过一些必要的改形后由定理 2 获得.

定理 2a. 如果条件 Ia, II 和 III 满足，存在连续算子 $\overline{K}^{-1}$，方程(1)具有解 x^*，则成立如下估计式

$$\|x^*-\varphi_0^{-1}\bar{x}^*\|\leqslant\bar{p}\|x^*\|, \tag{54}$$

其中

$$\bar{p}=(1+\varepsilon)|\lambda|\bar{\eta}\|\varphi_0^{-1}\overline{K}^{-1}\|+\varepsilon(1+\|\varphi_0^{-1}\overline{K}^{-1}\varphi K\|), \tag{55}$$

而

$$\varepsilon\leqslant\eta_1|\lambda|+\eta_2\|K\|.$$

作为 $\bar{p}$ 也可取

$$\bar{p}=2|\lambda|\bar{\eta}\|\varphi_0^{-1}\|\|\varphi_0^{-1}\overline{K}^{-1}\varphi_0\|+\varepsilon(1+\|\varphi_0^{-1}\overline{K}^{-1}\varphi K\|). \tag{56}$$

能用(56)式作为 $\bar{p}$ 的值这一点可从(13)式立即推知. 至于(55)式，则需要稍作补充论证. 保持定理 2 证明中所用的记号，由条件 Ia 有

$$\|\tilde{x}-\tilde{x}_0\|=\|\varphi_0^{-1}\overline{K}^{-1}\overline{K}\varphi_0\tilde{x}-\varphi_0^{-1}\overline{K}^{-1}\varphi_0\varphi_0^{-1}\varphi K\tilde{x}\|$$

$$\leqslant\|\varphi_0^{-1}\overline{K}^{-1}\|\|\overline{K}\varphi_0\tilde{x}-\varphi K\tilde{x}\|\leqslant|\lambda|\bar{\eta}\|\varphi_0^{-1}\overline{K}^{-1}\|\|\tilde{x}\|.$$

利用此估计式代替(17)，即得到(55).

注意，在定理 2 的注中所讲到的那个结果，只要将叙述的形式作相应的改变则仍然成立.

最后，我们叙述一个关于近似解序列的收敛性定理.

定理 3a. 如果下述条件满足:

1) 算子 K 有连续逆算子;

2) 对每个 $n=1, 2, \cdots$ 条件 Ia, II, III 成立，且

$$\lim_{n\to\infty}\bar{\eta}\|\varphi_0^{-1}\|=\lim_{n\to\infty}\eta_1\|\varphi_0^{-1}\varphi\|=\lim_{n\to\infty}\eta_2\|\varphi_0^{-1}\varphi\|=0. \tag{57}$$

那么，当 n 充分大时近似方程(48)可解且近似解序列收敛于精确解. 这时

$$\|x^*-\tilde{x}_n^*\|\leqslant\bar{Q}\bar{\eta}\|\varphi_0^{-1}\|+\bar{Q}_1\eta_1\|\varphi_0^{-1}\varphi\|+\bar{Q}_2\eta_2\|\varphi_0^{-1}\varphi\|, \tag{58}$$

其中 $\tilde{x}_n^*=\varphi_0^{-1}x_n^*$ 而 $\bar{Q}$, $\bar{Q}_1$ 和 $\bar{Q}_2$ 是常数.

事实上，由(50)和(46)可知定理 3 的全部条件均满足，由此即得上述的结果.

对于定理 4 的完全叙述就不再进行了. 我们指出，条件(25)和(26)在新的表示下分别是

$$\bar{r}=|\lambda|[(1+|\lambda|\eta_1)\bar{\eta}\|\varphi_0^{-1}\bar{K}^{-1}\|+\eta_1\|\varphi_0^{-1}\bar{K}^{-1}\varphi K\|]<1 \tag{59}$$

和

$$\|K^{-1}\|\leqslant\frac{1+\|\varphi_0^{-1}\bar{K}^{-1}\varphi\|+|\lambda|\bar{\eta}\|\varphi_0^{-1}\bar{K}^{-1}\|+\|\varphi_0^{-1}\bar{K}^{-1}\varphi K\|}{1-\bar{r}}. \tag{60}$$

为了证明这一点，应该进行如同对定理 4 所作的论证，利用定理 2a 的以(55)式为 $\bar{p}$ 值的估计式(54).

直接代入(25)和(26)得到 $\bar{r}$ 的值为

$$\begin{aligned}\bar{r}=|\lambda|[&(1+|\lambda|\eta_1)\bar{\eta}\|\varphi_0^{-1}\|\|\varphi_0^{-1}\bar{K}^{-1}\varphi_0\|\\&+\eta_1(1+\|\varphi_0^{-1}\bar{K}^{-1}\varphi K\|)],\end{aligned} \tag{61}$$

代替(60)式而得到如下的估计

$$\|K^{-1}\|\leqslant$$

$$\frac{1+\|\varphi_0^{-1}\bar{K}^{-1}\varphi\|+|\lambda|\bar{\eta}\|\varphi_0^{-1}\|\|\varphi_0^{-1}\bar{K}^{-1}\varphi_0\|+\|\varphi_0^{-1}\bar{K}^{-1}\varphi K\|}{1-\bar{r}}, \tag{62}$$

其中的 $\bar{r}$ 由(61)式给出.

§2. 可化为第二类方程的方程

2.1. 我们来考察这样的方程，它的左端虽然不能分出恒等算子，甚至它的左端的连续线性算子不是从给定空间到自身而是到另一个赋范空间，但由于这类算子的主要部分具有简单的形式，这类方程可以归结为我们在 § 1 中所讨论过的第二类方程.

于是，我们假定有两个赋范空间$\mathbf{X}$和$\mathbf{Y}$而且这两个空间都分离出一个完备的子空间$\tilde{\mathbf{X}}\subset\mathbf{X}$ 和 $\tilde{\mathbf{Y}}\subset\mathbf{Y}$. 又设 Φ 是从$\mathbf{Y}$到 $\tilde{\mathbf{Y}}$ 上的连续线性算子.

考察两个方程,即精确方程

$$K_1x\equiv Gx-\lambda Tx=y_1 \tag{1}$$

和与之对应的近似方程

$$\tilde{K}_1\tilde{x}\equiv G\tilde{x}-\lambda\tilde{T}\tilde{x}=\Phi y_1. \tag{2}$$

这里的G和T(从而算子K_1)是由$\mathbf{X}$到$\mathbf{Y}$的连续线性算子，算子$\tilde{T}$(从而算子$\tilde{K}_1$)是由$\tilde{\mathbf{X}}$到$\tilde{\mathbf{Y}}$的连续线性算子. 此外，对上述算子我们还作如下假定:

1) G具有连续逆算子;

2) 算子G建立了$\tilde{\mathbf{X}}$与$\tilde{\mathbf{Y}}$之间的一一对应,即$G(\tilde{\mathbf{X}})=\tilde{\mathbf{Y}}$,因而$G^{-1}(\tilde{\mathbf{Y}})=\tilde{\mathbf{X}}$.

我们再引进类似于 § 1 中条件 I—III 的条件.

Ib. 对任何 $\tilde{x}\in\tilde{\mathbf{X}}$,不等式

$$\|\Phi T\tilde{x}-\tilde{T}\tilde{x}\|\leqslant\mu\|\tilde{x}\|$$

成立.

IIb. 对于任何一个 $x\in\mathbf{X}$ 可求得 $\tilde{y}\in\tilde{\mathbf{Y}}$ 使得

$$\|Tx-\tilde{y}\|\leqslant\mu_1\|x\|.$$

IIIb. 存在元素 $\tilde{y}_1\in\tilde{\mathbf{Y}}$ 使得

$$\|y_1-\tilde{y}_1\|\leqslant\mu_2\|y_1\|.$$

现在来证明,在上述假定之下,对方程(1)和(2)的研究可归结为研究第二类方程.

事实上,对方程(1)和(2)的两端作用以算子G^{-1}可得

$$Kx\equiv G^{-1}K_1x\equiv x-\lambda G^{-1}Tx=G^{-1}y_1, \tag{3}$$

$$\tilde{K}\tilde{x}\equiv G^{-1}\tilde{K}_1\tilde{x}\equiv\tilde{x}-\lambda G^{-1}\tilde{T}\tilde{x}=G^{-1}\Phi y_1. \tag{4}$$

方程(3)和(4)具有 § 1 中方程(1)和(2)的形式，这时算子

$H, \widetilde{H}, P(K, \widetilde{K})$的作用分别由算子 $G^{-1}T$, $G^{-1}\widetilde{T}$, $G^{-1}\Phi G(G^{-1}K_1$, $G_1^{-1}\widetilde{K}_1)$担当, y 的作用由 $G^{-1}y_1$ 来担当.

现验证条件 I、II 和 III 均满足.

I. 由于 Ib, 则

$$\|PH\tilde{x}-\widetilde{H}\tilde{x}\|=\|G^{-1}\Phi GG^{-1}T\tilde{x}-G^{-1}\widetilde{T}\tilde{x}\|\leqslant\|G^{-1}\|\mu\|\tilde{x}\|.$$

II. 由于 IIb, 对 $x\in\mathbf{X}$, 我们求得相应的 $\tilde{y}\in\widetilde{\mathbf{Y}}$. 命 $\tilde{x}=G^{-1}\tilde{y}$, 则

$$\|Hx-\tilde{x}\|=\|G^{-1}Tx-G^{-1}\tilde{y}\|\leqslant\|G^{-1}\|\mu_1\|x\|.$$

III. 设 $\tilde{y}_1$ 是按 IIIb 所求得的元素, 命 $\tilde{y}=G^{-1}\tilde{y}_1$, 则

$$\|y-\tilde{y}\|=\|G^{-1}y_1-G^{-1}\tilde{y}_1\|\leqslant\|G^{-1}\|\mu_2\|y_1\|\leqslant\|G^{-1}\|\|G\|\mu_2\|y\|.$$

这样, §1 中的条件 I, II, III 均满足并且常数 η_1, η 和 η_2 通过 μ, μ_1 和 μ_2 表示如下:

$$\eta=\mu\|G^{-1}\|, \eta_1=\mu_1\|G^{-1}\|, \quad \eta_2=\mu_2\|G^{-1}\|\|G\|. \tag{5}$$

注. 代替 IIb 和 IIIb, 可以要求有满足不等式

$$\|G^{-1}Tx-\tilde{x}\|\leqslant\mu_1'\|x\|, \quad \|G^{-1}y_1-\tilde{y}\|\leqslant\mu_2'\|G^{-1}y_1\| \tag{6}$$

的元素 $\tilde{x}\in\widetilde{\mathbf{X}}$ 和 $\tilde{y}\in\widetilde{\mathbf{Y}}$ 存在. 这时 II 和 III 在 $\eta_1=\mu_1'$ 和 $\eta_2=\mu_2'$ 时成立.

由于对方程(3)和(4)来说条件 I, II 和 III 均满足, 所以可由上节的定理得到与之对应的关于方程(3)和(4)或与之等价的方程(1)和(2)的定理. 我们将要对相应的定理作全面的叙述, 但在此之前还要有一个进一步简化的命题.

亦即, 根据对应的映射 G, 我们确定空间 $\mathbf{X}$ 和 $\mathbf{Y}$ 中的范数为

$$\|x\|_{\mathbf{X}}=\|Gx\|_{\mathbf{Y}}; \quad \|y\|_{\mathbf{Y}}=\|G^{-1}y\|_{\mathbf{X}} \qquad (x\in\mathbf{X}, y\in\mathbf{Y}). \tag{7}$$

这时, G 显然是等距映射: $\|G\|=\|G^{-1}\|=1$. 因此

$$\begin{gathered}\|P\|=\|\Phi\|, \|K\|=\|K_1\|, \|\widetilde{K}\|=\|\widetilde{K}_1\|,\\ \|H\|=\|T\|, \quad \|\widetilde{H}\|=\|\widetilde{T}\|.\end{gathered} \tag{8}$$

2.2. 现在对方程(1)和(2)叙述 §1 中的基本定理.

定理 1b. 如果条件 Ib 和 IIb 满足, 存在连续算子 K_1^{-1} 且

$$q=|\lambda|[\mu+\|I-\Phi\|\mu_1]\|K_1^{-1}\|<1, \tag{9}$$

则算子 $\widetilde{K}_1$ 也存在连续逆算子，并且

$$\|\widetilde{K}_1^{-1}\|\leqslant\frac{\|K_1^{-1}\|}{1-q}. \tag{10}$$

定理 2b. 如果条件 Ib, IIb, IIIb满足，存在连续算子 $\widetilde{K}_1^{-1}$(特别，如果定理 1b 的条件成立)，并且方程(1)的解 x^*存在，则有如下估计

$$\|x^*-\tilde{x}^*\|\leqslant p\|x^*\|, \tag{11}$$

其中 $\tilde{x}^*$ 是方程(2)的解，而

$$p=2|\lambda|\mu\|\widetilde{K}_1^{-1}\|+(\mu_1|\lambda|+\mu_2\|K_1\|)(1+\|\widetilde{K}_1^{-1}\Phi K_1\|). \tag{12}$$

注. 如果存在这样的 $\tilde{x}\in\mathbf{X}$ 使得

$$\|x^*-\tilde{x}\|\leqslant\varepsilon\|x^*\|, \tag{13}$$

则估计式(11)在不要求条件 IIb 和 IIIb 满足的情况下仍然成立，此时 p 可取如下值:

$$p=2|\lambda|\mu\|\widetilde{K}_1^{-1}\|+\varepsilon(1+\|\widetilde{K}_1^{-1}\Phi K_1\|). \tag{14}$$

定理 3b. 如果对每个 $n=1,2,\cdots$条件 Ib, IIb 和IIIb均满足，算子 K_1 有连续逆，并且成立如下关系

$$\lim_{n\to\infty}\mu=\lim_{n\to\infty}\mu_1\|\Phi\|=\lim_{n\to\infty}\mu_2\|\Phi\|=0, \tag{15}$$

则近似解(亦即方程(2)的解)序列收敛于方程(1)的解x^*. 在这种情形下

$$\|x^*-\tilde{x}_n^*\|\leqslant Q\mu+Q_1\mu_1\|\Phi\|+Q_2\mu_2\|\Phi\|. \tag{16}$$

定理 4b. 如果存在连续算子 $\widetilde{K}_1^{-1}$, 条件 Ib 和 IIb 满足且

$$r=|\lambda|\mu(1+|\lambda|\mu_1)\|\widetilde{K}_1^{-1}\|+|\lambda|\mu_1(1+\|\widetilde{K}_1^{-1}\Phi K_1\|)<1, \tag{17}$$

则 $\tilde{K}_1$ 存在连续左逆，且

$$\|K_1^{-1}\| \leqslant \frac{1+\|\widetilde{K}_1^{-1}\Phi\|+|\lambda|\mu\|\widetilde{K}_1^{-1}\|+\|\widetilde{K}_1^{-1}\Phi K_1\|}{1-r}. \tag{18}$$

2.3. 最后，我们指出一种特殊情形，这时定理的叙述可以得到简化．也就是说，此时近似方程(2)是通过将精确方程进行投影这种特殊方法构成的．亦即在方程(1)的两端作用算子Φ而得到近似方程

$$\widetilde{K}_1\tilde{x}=\Phi K_1\tilde{x}\equiv G\tilde{x}-\lambda\Phi T\tilde{x}=\Phi y_1, \tag{19}$$

这时 $\tilde{T}=\Phi T$.

易见，此时在 $\mu=0$ 时条件 Ib成立．事实上，

$$\|\tilde{T}\tilde{x}-\Phi T\tilde{x}\|=\|\Phi T\tilde{x}-\Phi T\tilde{x}\|=0.$$

因此，在叙述定理时可取 $\mu=0$.

我们再补充一个从定理 1.6 得出的并与所考察的方程类型有关的定理.

定理 6b. 如果存在形如(19)的近似方程序列及与之相应的映射序列 Φ_n，并且

1) 空间 $\mathbf{Y}$ 是完备的；

2) 在 $\mathbf{Y}$ 上有 $\lim\limits_{n\to\infty}\Phi_n=I$;

3) 算子 $G^{-1}T$ 是紧算子，

还有，如果存在 K_1^{-1}，则对于充分大的 n，近似方程可解并且近似解序列收敛于精确解.

§3. 对无限方程组的应用*)

3.1. 设给定无限方程组

*) 关于无限方程组，参见 Riesz [4]，Hellinger 和 Toeplitz[1]，也可参见 Канторович 和 Крылов.

$$\xi_j-\lambda\sum_{k=1}^{\infty}a_{jk}\xi_k=b_j \qquad (j=1,2,\cdots,) \tag{1}$$

并假定

$$\sum_{j,k=1}^{\infty}|a_{jk}|^2<\infty, \quad \sum_{j=1}^{\infty}|b_j|^2<\infty. \tag{2}$$

我们来研究这个方程组的满足条件

$$\sum_{k=1}^{\infty}|\xi_k|^2<\infty \tag{3}$$

的解.

普遍采用的方法之一就是所谓的简化法. 这种解法是用 n 个自变量 n 个方程的方程组

$$\xi_j-\lambda\sum_{k=1}^{n}a_{jk}\xi_k=b_j \qquad (j=1,2,\cdots,n) \tag{4}$$

来代替无限方程组(1), 而将(4)的解作为(1)的近似解.

我们感兴趣的是近似解的误差估计以及当 $n\to\infty$ 时近似解向精确解的收敛问题.

为了应用一般理论, 取 $\mathbf{X}=\mathbf{l}^2$. 此时, 方程组(1)可写成 $\mathbf{X}$ 中一个方程的形式

$$Kx\equiv x-\lambda Hx=y \qquad (x=\{\xi_n\},\ y=\{b_n\}). \tag{5}$$

这里的H应理解成由方程组

$$z=Hx, \quad \zeta_j=\sum_{k=1}^{\infty}a_{jk}\xi_k$$

$$(j=1,2,\cdots;\ x=\{\xi_n\},\ z=\{\zeta_n\})$$

所确定的作用于 $\mathbf{l}^2$ 中的连续线性算子(因而也是紧算子, 参见 **XI**. 2.2).

作为 $\overline{\mathbf{X}}$, 我们取有限维欧氏空间 $\mathbf{l}_n^2$. 这时, 自然是取 $\mathbf{l}^2$ 中从第$(n+1)$个坐标起全为零的元素全体作为 $\tilde{\mathbf{X}}$. 算子 φ_0 和 φ_0^{-1}

的意义不用说明也是清楚的. φ 是将元素 $x=\{\xi_m\}\in \mathbf{l}^2$ 和元素

$$\bar{x}=\varphi x=(\xi_1, \xi_2, \cdots, \xi_n)\in \mathbf{l}_n{}^2$$

对应起来的算子. 显然

$$\|\varphi\|=\|\varphi_0\|=\|\varphi_0^{-1}\|=1.$$

方程组(4)现在可以写成 $\overline{\mathbf{X}}$ 中一个方程的形式

$$\overline{K}\bar{x}\equiv\bar{x}-\lambda\overline{H}\bar{x}=\varphi y, \tag{6}$$

其中算子 $\overline{H}$ 由"截断"矩阵

$$A_n=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

来确定.

我们现在来验证条件 Ia, II 和 III 均满足. 对任何 $\tilde{x}=(\xi_1, \xi_2, \cdots, \xi_n, 0, \cdots)\in\widetilde{\mathbf{X}}$, 有

$$\bar{z}=\varphi H\tilde{x}-\overline{H}\varphi_0\tilde{x} \qquad (\bar{z}=(\zeta_1, \zeta_2, \cdots, \zeta_n)),$$

$$\zeta_j=\sum_{k=1}^{n}a_{jk}\xi_k-\sum_{k=1}^{n}a_{jk}\xi_k=0 \quad (j=1, 2, \cdots, n).$$

因此条件 Ia, 在 $\bar{\eta}=0$ 时成立.

为了验证条件 II, 取任意的 $x=\{\xi_m\}\in \mathbf{l}^2$ 并设 $\tilde{x}=[Hx]_n=(\zeta_1, \zeta_2, \cdots, \zeta_n, 0, \cdots)$*). 这时

$$\|Hx-\tilde{x}\|=\left[\sum_{j=n+1}^{\infty}\sum_{k=1}^{\infty}|a_{jk}\xi_k|^2\right]^{\frac{1}{2}}$$

$$\leqslant\left[\sum_{j=n+1}^{\infty}\sum_{k=1}^{\infty}|a_{jk}|^2\sum_{k=1}^{\infty}|\xi_k|^2\right]^{\frac{1}{2}}=$$

---

*) 符号 $[x]_n$ 表示"截断"元素, 也就是由 $x\in \mathbf{l}^2$ 得到的从第 $(n+1)$ 个起所有坐标均为零的元素. 显然 $[x]_n=\varphi_0^{-1}\varphi x=Px$.

$$=\eta_1\|x\|,$$

其中

$$\eta_1=\left[\sum_{j=n+1}^{\infty}\sum_{k=1}^{\infty}|a_{jk}|^2\right]^{\frac{1}{2}}.$$

显然，由(2)可知当 $n\to\infty$ 时 $\eta_1\to 0$.

最后，令 $\tilde{y}=[y]_n$，则

$$\|\tilde{y}-y\|=\left[\sum_{j=n+1}^{\infty}|b_j|^2\right]^{\frac{1}{2}}$$

$$=\left[\frac{\sum\limits_{j=n+1}^{\infty}|b_j|^2}{\sum\limits_{j=1}^{\infty}|b_j|^2}\right]^{\frac{1}{2}}\|y\|,$$

从而，在条件 III 中可取

$$\eta_2=\left[\frac{\sum\limits_{j=n+1}^{\infty}|b_j|^2}{\sum\limits_{j=1}^{\infty}|b_j|^2}\right]^{\frac{1}{2}},$$

这里当 $n\to\infty$ 时 $\eta_2\to 0$.

因此，这时可以应用在 I. 9 中所述的定理. 特别，定理 3a 使我们可以断言：如果 λ 不是(1)的特征值(也就是说若 λ 不是算子 H 的特征值)，则当 n 充分大时方程组(4)可解，而且近似解序列收敛于精确解. 收敛速度由下列不等式决定：

$$\|x^*-\varphi_0^{-1}\bar{x}_n{}^*\|$$

$$\leqslant Q_1\left[\sum_{j=n+1}^{\infty}\sum_{k=1}^{\infty}|a_{jk}|^2\right]^{\frac{1}{2}}+Q_2\left[\frac{\sum\limits_{j=n+1}^{\infty}|b_j|^2}{\sum\limits_{j=1}^{\infty}|b_j|^2}\right]^{\frac{1}{2}},$$

其中 $x^*=(\xi_1^*, \xi_2^*, \cdots, \xi_n^*, \cdots)$ 表示(1)的解而 $\bar{x}^*=(\xi_1^{(n)}, \xi_2^{(n)}, \cdots, \xi_n^{(n)})$表示(4)的解.

由此易见,对 $k=1, 2, \cdots, n$, ξ_k^* 和 $\xi_k^{(n)}$的差别很小,而当 $k>n$ 时坐标 ξ_k^* 很小. 由此还显然可知有

$$\lim_{n\to\infty}\xi_k^{(n)}=\xi_k^* \qquad (k=1, 2, \cdots).$$

利用定理 1.4(具有 §1 中的条件(59))可得到下面的定理.

定理 1: 如果方程组(4)有唯一解, 且

$$\bar{r}=|\lambda|\eta_1\|\bar{K}^{-1}\|\|\varphi K\|<1, \tag{7}$$

则 λ 不是方程组(1)的特征值, 而且

$$\|K^{-1}\|\leqslant\frac{1+\|\bar{K}^{-1}\|(1+\|\varphi K\|)}{1-\bar{r}}.$$

这个定理值得注意的是它建立了根据有限方程组的可解性性态而断言相应的无限方程组可解性的关系,而且不等式(7)中所出现的量均不难求得,因而本定理所给出的可解性的判别准则是十分有效的.

量 η_1 的值上面已经指出过了. 对于$\|\varphi K\|$可用下估计式

$$\|\varphi K\|\leqslant 1+\|\varphi H\|\leqslant 1+\left[\sum_{j=1}^{n}\sum_{k=1}^{\infty}|a_{jk}|^2\right]^{\frac{1}{2}}.$$

最后,如果 A_n 是对称矩阵,则$\|\bar{K}^{-1}\|$由下式

$$\|\bar{K}^{-1}\|=\max_{j=1,2,\cdots,n}\frac{1}{|1-\lambda\lambda_j^{(n)}|}$$

确定,其中 $\lambda_j^{(n)}$ 是矩阵 A_n 的特征值. 在一般情形下, $\|\bar{K}^{-1}\|$的表达式比较复杂(见 V. 2.8).

注 1: 代替满足条件(2)的方程组,我们还可以考察更一般的方程组类, 即考察当 H 是由 $\mathbf{l}^2$ 到 $\mathbf{l}^2$ 内的紧算子的情形. 这时 η_1 由下式

$$\eta_1=\|H-\varphi_0^{-1}\varphi H\varphi_0^{-1}\varphi\|$$

确定，和前面一样，当 $n\to\infty$ 时 $\eta_1\to 0$(XI. 2. 2)，从而这时可以应用一般理论中的所有定理.

注 2: 一般理论还可应用到其他一些类型的无限方程组上去，即可将这些方程组看作是在别的序列空间中的泛函方程。例如，满足条件

$$\sum_{j=1}^{\infty}\left[\sum_{k=1}^{\infty}|a_{jk}|^{p/p-1}\right]^{p-1}<\infty \qquad (1<p<\infty)$$

的 Riesz 型方程组(这时，H 应视为 l^p 中的算子)就是这类方程组. 正则和完全正则方程组也是这一类方程组(在空间 l^∞ 中).

§4. 在积分方程中的应用

4.1. 积分方程数值解的最有效的方法之一就是借助于求积公式用线代数方程组去代替积分方程.

设给出积分方程

$$x(s)-\lambda\int_0^1 h(s,t)x(t)dt=y(s). \tag{1}$$

采用按结点 t_1, t_2, $\cdots$, t_n 构造的并具有如下形式的数值求积公式

$$\int_0^1 x(t)dt=\sum_{k=1}^{n}A_k x(t_k), \tag{2}$$

而且仅要求在诸结点处等式(1)满足. 于是，我们即得到方程组

$$x(t_j)-\lambda\sum_{k=1}^{n}A_k h(t_j,t_k)x(t_k)=y(t_j) \qquad (j=1,2,\cdots,n), \tag{3}$$

这个方程组的解就是所求的积分方程的解在点 $t_1, t_2, \cdots, t_n$ 上的近似值.

我们现在来讨论一下如何应用一般理论来研究这个方法的误差估计. 为确定起见,假定方程(1)的右端是以 1 为周期的连续函数,核函数 $h(s,t)$既是关于 s 的又是关于 t 的以 1 为周期的连续函数. 在这些条件之下,方程(1)的解也将是连续的周期函数. 我们将把方程(1)看作连续周期函数空间 $\mathbf{X}=\tilde{\mathbf{C}}$ 中的泛函方程. 代数方程组(3)看作是空间 $\tilde{\mathbf{X}}=\mathbf{l}_n^\infty$ 中的近似泛函方程

$$\bar{K}\bar{x}\equiv\bar{x}-\lambda\bar{H}\bar{x}=\varphi y. \tag{4}$$

此外,求积公式(2)的结点和系数取为

$$A_k=\frac{1}{n},\ t_k=\frac{2k-1}{2n}\qquad(k=1,2,\cdots,n),$$

也就是说采用平均矩形公式,这对于周期函数是最方便的一种方法. 在这种选择之下,方程(4)中的算子 $\bar{H}$ 由下述矩阵

$$\begin{pmatrix}\frac{1}{n}h\left(\frac{1}{2n},\frac{1}{2n}\right) & \frac{1}{n}h\left(\frac{1}{2n},\frac{3}{2n}\right) & \cdots & \frac{1}{n}h\left(\frac{1}{2n},\frac{2n-1}{2n}\right)\\ \frac{1}{n}h\left(\frac{3}{2n},\frac{1}{2n}\right) & \frac{1}{n}h\left(\frac{3}{2n},\frac{3}{2n}\right) & \cdots & \frac{1}{n}h\left(\frac{3}{2n},\frac{2n-1}{2n}\right)\\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \frac{1}{n}h\left(\frac{2n-1}{2n},\frac{1}{2n}\right) & \frac{1}{n}h\left(\frac{2n-1}{2n},\frac{3}{2n}\right) & \cdots & \frac{1}{n}h\left(\frac{2n-1}{2n},\frac{2n-1}{2n}\right)\end{pmatrix} \tag{5}$$

确定. 算子 φ 按以下方式

$$\varphi x=(x(t_1),x(t_2),\cdots,x(t_n)),\ \|\varphi\|=1\quad(x\in\tilde{\mathbf{C}}).$$

确定. 至于空间 $\tilde{\mathbf{X}}$ 和与之相联系的映射 φ_0^{-1}, 则我们将用两种不同的方法来确定,从而也就得到两种误差估计.

首先, 取所有在区间$[t_k,t_{k+1}]$上为线性的连续周期函数的集合为$\tilde{\mathbf{X}}$, 其中 $t_k=\frac{2k-1}{2n}(k=0,\pm1,\pm2,\cdots)$. 每个这样的函数是由在点 $t_1,t_2,\cdots,t_n$ 处的取值来确定的, 从而映射 φ_0^{-1} 这样来定义:

$$\tilde{x}=\varphi_0^{-1}\bar{x},$$

$(\tilde{x}=(\xi_1,\xi_2,\cdots,\xi_n)\in \mathbf{l}_n^\infty;\ \tilde{x}(t_k)=\xi_k;\ k=1,2,\cdots,n;\ \tilde{x}\in\tilde{\mathbf{X}})$. 这时，$\|\varphi_0^{-1}\|=1$.

用 $\omega_s(\delta)$ 表示函数 $h(s,t)$ 的连续模，它作为 s 的函数定义如下:

$\omega_s(\delta)=\sup|h(s+\sigma,t)-h(s,t)|\quad (0\leqslant s,\ t\leqslant 1;\ |\sigma|\leqslant\delta)$, $\omega_t(\delta)$的定义类似.

为了验证条件 Ia, 我们对形如 $z(t)\tilde{x}(t)$ 的函数的求积公式(2)作误差估计，这里的 $z(t)$是周期函数而且 $z(t)$的连续模不超过 $\omega(\delta)$, 而 $\tilde{x}\in\tilde{\mathbf{X}}$. 由周期性我们有

$$\begin{aligned}
&\left|\int_0^1 z(t)\tilde{x}(t)dt-\sum_{k=1}^{n}\frac{1}{n}z(t_k)\tilde{x}(t_k)\right|\\
&=\left|\int_{t_0}^{t_n} z(t)\tilde{x}(t)dt-\sum_{k=0}^{n-1}\frac{1}{2n}(z(t_k)\tilde{x}(t_k)\right.\\
&\qquad\left.+z(t_{k+1})\tilde{x}(t_{k+1}))\right|\\
&=\left|\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\left[z(t)-z\left(\frac{t_k+t_{k+1}}{2}\right)\right]\tilde{x}(t)dt\right.\\
&\quad+\sum_{k=0}^{n}\frac{1}{2n}\left[\tilde{x}(t_k)\left(z\left(\frac{t_k+t_{k+1}}{2}\right)-z(t_k)\right)\right.\\
&\quad\left.\left.+\tilde{x}(t_{k+1})\left(z\left(\frac{t_k+t_{k+1}}{2}\right)-z(t_{k+1})\right)\right]\right|\\
&\leqslant 2\,\omega\left(\frac{1}{2n}\right)\|\tilde{x}\|,
\end{aligned}$$

这里，由于函数 $\tilde{x}(t)$的线性性，故在每个区间中

$$\int_{t_k}^{t_{k+1}}\tilde{x}(t)dt=\frac{1}{2n}(\tilde{x}(t_k)+\tilde{x}(t_{k+1}))\qquad (k=0,1,\cdots,n-1).$$

在前面的不等式中命 $z(t)=h(t_j,t)$, 得到

$$\|\varphi H\tilde{x}-\bar{H}\varphi_0\tilde{x}\|=\max_{j=1,2,\cdots,n}\left|\int_0^1 h(t_j,t)\tilde{x}(t)dt-\right.$$

$$-\frac{1}{n}\sum_{k=1}^{n}h(t_j, t_k)\tilde{x}(t_k)\Big| \leqslant 2\omega_t\left(\frac{1}{2n}\right)\|\tilde{x}\|.$$

于是，条件 Ia 在 $\bar{\eta}=2\omega_t\left(\frac{1}{2n}\right)$时成立.

现在来验证条件 II. 我们先来证明，如果 $z\in\tilde{\mathbf{C}}$ 是连续模不超过 $\omega(\delta)$的任意函数，则有

$$\|z-\tilde{z}\|\leqslant\omega\left(\frac{1}{n}\right),$$

其中 $\tilde{z}=\varphi_0^{-1}\varphi z$,即 $\tilde{z}$ 是分段线性函数，它在点 $t_1, t_2, \cdots, t_n$处的值和函数 $z(t)$ 在这些点上的对应值相等. 事实上，如果 $t_j\leqslant s\leqslant t_{j+1}$，则

$$\begin{aligned}|z(s)-\tilde{z}(s)| &= |z(s)-[(t_{j+1}-s)z(t_j)\\ &\quad +(s-t_j)z(t_{j+1})]n|\\ &\leqslant n[|(t_{j+1}-s)(z(s)-z(t_j))|+|(s-t_j)(z(s)\\ &\quad -z(t_{j+1}))|]\\ &\leqslant n\omega\left(\frac{1}{n}\right)(t_{j+1}-t_j)=\omega\left(\frac{1}{n}\right),\end{aligned}$$

于是，由 s 的任意性即得到要证的结果，现在来估计元素 $z=Hx$ $(x\in\tilde{\mathbf{C}})$的连续模:

$$\begin{aligned}|z(s)-z(s')|&\leqslant\int_0^1|h(s,t)-h(s',t)||x(t)|dt\\ &\leqslant\omega_s(\delta)\|x\| \qquad (|s-s'|\leqslant\delta).\end{aligned}$$

于是，作为 $\omega(\delta)$可取

$$\omega(\delta)=\omega_s(\delta)\|x\|,$$

由以上所证，如果命 $\tilde{z}=\varphi_0^{-1}\varphi Hx$，则

$$\|Hx-\tilde{z}\|\leqslant\omega_s\left(\frac{1}{n}\right)\|x\|.$$

因而，条件 II 在

$$\eta_1=\omega_s\left(\frac{1}{n}\right)$$

时成立．最后，将上述讨论应用于元素 y，求得 $\tilde{y}\in\tilde{\mathbf{X}}$ 使得

$$\|y-\tilde{y}\|\leqslant\varpi\left(\frac{1}{n}\right),$$

其中 ϖ 是函数 y 的连续模．由此可见条件 III 在

$$\eta_2=\frac{1}{\|y\|}\varpi\left(\frac{1}{n}\right)$$

时成立．

由核函数 $h(s,\ t)$ 和方程(1)右端 $y(s)$ 的连续性可知当 $n\to\infty$ 时 $\bar{\eta},\eta_1,\eta_2$ 均收敛于零，而 $\|\varphi\|=\|\varphi_0^{-1}\|=1$，因而可以应用一般理论的定理．特别，如果 λ 不是方程(1)的特征值，则由定理 1.3a 可知对于充分大的 n，方程组(3)可解而且近似解 $\varphi_0^{-1}\bar{x}_n^*$（是按由方程组(3)所求得的值 $\{\xi_k\}$ 构成的分段线性函数）收敛于精确解．

注意，由定理 1.1a，逆算子的范数 $\|\bar{K}^{-1}\|$ 有界而且不依赖于 n．指出这一点对以后是重要的．

如果 $h(s,t)$ 和 $y(s)$ 满足指数为 α 的 Lipschitz 条件，则由定理 1.3a 可知近似解序列的收敛速度可具有如下的估计

$$\|x^*-\varphi_0^{-1}\bar{x}_n^*\|=O\left(\frac{1}{n^\alpha}\right).$$

事实上，这时 $\omega_s(\delta)$，$\omega_t(\delta)$ 和 $\varpi(\delta)$ 具有形式 $O(\delta^\alpha)$．

4.2. 现在，我们采用另外的办法选择子空间 $\tilde{\mathbf{X}}$，再来讨论这种办法的误差估计的方法问题．

为方便起见，我们可认为 n 是奇数：$n=2m+1$．取所有次数不超过 m 的三角多项式的集合为 $\tilde{\mathbf{X}}$，也就是说 $\tilde{x}\in\tilde{\mathbf{X}}$ 乃表示 $\tilde{x}$ 可表达为如下的三角多项式

$$\tilde{x}(t)=\frac{a_0}{2}+\sum_{k=1}^{m}(a_k\cos 2\pi kt+b_k\sin 2\pi kt).$$

易见，算子 φ_0^{-1} 将元素 $\bar{x}=(\xi_1, \xi_2, \cdots, \xi_n)\in \mathbf{l}_n^\infty$ 和在结点 $t_1, t_2, \cdots, t_n$ 处分别取值为 $\xi_1, \xi_2, \cdots, \xi_n$ 的内插多项式对应起来. 由 Бернштейн 引理*)，如果对三角多项式 $\tilde{x}$ 在结点 $t_1, t_2, \cdots, t_n$ 处有估计式

$$|\tilde{x}(t_j)|\leqslant 1 \qquad (j=1, 2, \cdots, n),$$

则

$$|\tilde{x}(t)|\leqslant A\ln m+B \qquad (0\leqslant t\leqslant 1).$$

由此可得到算子 φ_0^{-1} 的范数估计

$$\|\varphi_0^{-1}\|\leqslant A\ln m+B. \tag{6}$$

其次，以 E_m^t(或 E_m^s)表示用 m 阶三角多项式对 $h(s, t)$ 作为 t (或 s)的函数的最佳逼近，换言之，我们将假定存在下列形式的函数

$$h_1(s, t)=\frac{a_0'(s)}{2}+\sum_{k=1}^{m}[a_k'(s)\cos 2\pi kt + b_k'(s)\sin 2\pi kt],$$

$$h_2(s, t)=\frac{a_0''(t)}{2}+\sum_{k=1}^{m}[a_k''(t)\cos 2\pi ks + b_k''(t)\sin 2\pi ks]$$

使得

$$|h(s, t)-h_1(s, t)|\leqslant E_m^t, \ |h(s, t)-h_2(s, t)|\leqslant E_m^s \quad (0\leqslant s, t\leqslant 1).$$

为了验证条件 Ia，我们首先指出，如果被积函数是次数不超过 $n-1=2m$ 的三角多项式，则求积公式(2)是精确成立的等式**).

由于 $h_1(s, t)\tilde{x}(t)$ 正是这样的 t 的多项式，所以有

*）见 Натансон-I 第 542 页.

**）见 Натансон-I 第 611 页.

$$\int_0^1 h_1(s,t)\tilde{x}(t)dt=\frac{1}{n}\sum_{k=1}^{n}h_1(s,t_k)\tilde{x}(t_k)\qquad(0\leqslant s\leqslant 1).$$

因此

$$\begin{aligned}
&\|\varphi Hx-\bar{H}\varphi\tilde{x}\|\\
&=\max_{j=1,2,\cdots,n}\left|\int_0^1 h(t_j,t)\tilde{x}(t)dt-\frac{1}{n}\sum_{k=1}^{n}h(t_j,t_k)\tilde{x}(t_k)\right|\\
&\leqslant\max_{j=1,2,\cdots,n}\left|\int_0^1[h(t_j,t)-h_1(t_j,t)]\tilde{x}(t)dt\right|\\
&\quad+\max_{j=1,2,\cdots,n}\frac{1}{n}\left|\sum_{k=1}^{n}[h(t_j,t_k)-h_1(t_j,t_k)]\tilde{x}(t_k)\right|\\
&\leqslant 2E_m^t\|\tilde{x}\|.
\end{aligned}$$

因此，条件 Ia 在 $\bar{\eta}=2E_m^t$ 时成立.

为了验证元素 $z=Hx$ 可逼近的条件 II，我们来构造函数*)

$$\tilde{x}(s)=\int_0^1 h_2(s,t)x(t)dt.$$

由 $h_2(s,t)$ 的形式可知 $\tilde{x}(s)$ 是阶数为 m 的三角多项式，亦即 $\tilde{x}\in\tilde{\mathbf{X}}$. 同时

$$\|Hx-\tilde{x}\|=\max_s\left|\int_0^1[h(s,t)-h_2(s,t)]x(t)dt\right|\leqslant E_m^s\|x\|.$$

因此条件 II 在 $\eta_1=E_m^s$ 时成立.

最后，设 $E_m(y)$ 是用 m 阶三角多项式对函数 y 的最佳逼近，并命

$$\eta_2=\frac{1}{\|y\|}E_m(y),$$

---

*) 不难指出，函数 $h_2(s,t)$ 总可选得关于 t 连续，于是积分 $\int_0^1 h_2(s,t)x(t)dt$ 存在；另外，如果此积分按照 Banach 意义来理解(见 II.4.2)，则不必讲明这一点，因为这时积分对任意的有界函数有定义.

则条件 III 也满足.

由这个估计,利用定理 1.2a 可得到精确解与近似解之间接近程度的如下估计

$$\|x^*-\varphi_0^{-1}\bar{x}^*\|\leqslant\bar{p}\|x^*\|,$$

这里

$$\bar{p}=4|\lambda|E_m^t\|\varphi_0^{-1}\|\|\bar{K}^{-1}\|$$
$$+\left(|\lambda|E_m^s+\frac{E_m(y)\|K\|}{\|y\|}\right)(1+\|\varphi_0^{-1}\|\|\bar{K}^{-1}\|\|K\|).$$

正如在 4.1 中已建立的那样,$\|\bar{K}^{-1}\|$不依赖于 n 而有界.

假定核函数$h(s,t)$具有直到 ν 阶的关于 s 和 t 的各阶导数并且$\frac{\partial^\nu h}{\partial s^\nu}$和$\frac{\partial^\nu h}{\partial t^\nu}$满足指数为$\alpha$的 Lipschitz 条件, 对于函数 $y(s)$也作同样的假定. 这时, 由著名的 Jackson 定理(参见 Натансон-I 第 121 页),有

$$E_m^t=O\left(\frac{1}{m^{\nu+\alpha}}\right)=O\left(\frac{1}{n^{\nu+\alpha}}\right),\qquad E_m^s=O\left(\frac{1}{m^{\nu+\alpha}}\right)=O\left(\frac{1}{n^{\nu+\alpha}}\right)$$

$$E_m(y)=O\left(\frac{1}{m^{\nu+\alpha}}\right)=O\left(\frac{1}{n^{\nu+\alpha}}\right),$$

又因为由(6)有

$$\|\varphi_0^{-1}\|=O(\ln m)=O(\ln n),$$

所以由定理 1.3a 我们就得到近似解误差下降速度的如下估计

$$\|x^*-\varphi_0^{-1}\bar{x}_n^*\|=O\left(\frac{\ln n}{n^{\nu+\alpha}}\right).$$

利用定理 1.4, 可根据近似解的结果来建立方程(1)特征值的可能分布区域.

最后,我们指出,可用类似的讨论来估计当用其他求积公式时此方法的误差. 特别,对于非周期函数的情形,利用 Gauss公式需要采用代数多项式作插值多项式并利用代数多项式逼近函数的阶

的定理.

4.3. 考察积分方程(1)的其他一些近似解法，首先是矩量法. 这个方法是要求方程(1)的形如

$$\tilde{x}(t)=\sum_{k=1}^{n}c_k\omega_k(t) \qquad (0\leqslant t\leqslant 1) \tag{7}$$

的近似解，其中 $\omega_k(k=1,2,\cdots)$是完全正交函数系. 这时，系数c_k由方程组

$$c_j-\lambda\sum_{k=1}^{n}c_k\int_0^1\int_0^1 h(s,t)\omega_j(s)\omega_k(t)dsdt$$
$$=\int_0^1 y(s)\omega_j(s)ds \qquad (j=1,2,\cdots,n) \tag{8}$$

来确定. 这个方程组把使差 $Kx-y$ 为零的要求改为对已知正交函数系的前 n 个函数和 $Kx-y$ 正交的条件.

假定核函数 $h(s,t)$满足条件

$$\int_0^1\int_0^1|h(s,t)|^2dsdt<\infty,$$

并且取 $\mathbf{X}$ 为空间$\mathbf{L}^2(0,1)$而取$\tilde{\mathbf{X}}$为形如(7)的元素构成的子空间. 算子P定义为 $\tilde{\mathbf{X}}$ 上的正交投影算子，亦即 $\tilde{x}=Px$ 表示

$$\tilde{x}=\sum_{k=1}^{n}(x,\omega_k)\omega_k.$$

方程组(8)可以写成空间 $\tilde{\mathbf{X}}$ 中一个方程的形式

$$\tilde{x}-PH\tilde{x}=Py.$$

由此可见，我们要得到的是在 1.8 中已经提到过的那些条件. 这时，由于函数系$\{\omega_k\}$是完全的，定理 1.6 的条件成立，因而近似解收敛于精确解. 收敛的速度取决于核函数 $h(s,t)$ 和右端函数 $y(s)$的性质，它依赖于这两个函数按正交函数系的正交展开式的

收敛速度. 我们来讨论当核函数 $h(s, t)$ 是周期函数而且关于 s 满足以 $\alpha+\frac{1}{2}(\alpha>0)$ 为指数的 Lipschitz 条件的情形. 对 $y(s)$ 也提出类似的要求. 取三角函数系为正交函数系 $\{\omega_k\}$. 这时由 Jackson 定理不难得到估计

$$\eta_1=O(n^{-1/2-\alpha}),\quad \eta_2=O(n^{-1/2-\alpha}),$$

因而

$$\|x^*-\tilde{x}_n^*\|_{L^2}=O(\eta_1+\eta_2)=O(n^{-1/2-\alpha}). \tag{9}$$

我们来证明,在这种情形下近似解是一致收敛于精确解的.

设 $\tilde{x}=\sum_{j=1}^{n}c_j\omega_j$. 我们用 $\|\tilde{x}\|_{L^2}$ 来估计 $\|\tilde{x}\|_C$. 因为 $\|\omega_j\|_C\leqslant M$ $(j=1,2,\cdots)$, 故利用 Cauchy-Буняковский 不等式得

$$\|\tilde{x}\|_C\leqslant M\sum_{k=1}^{n}|c_j|\leqslant M\sqrt{n}\left[\sum_{j=1}^{n}|c_j|^2\right]^{1/2}$$

$$=M\sqrt{n}\|x\|_{L^2}.$$

将这个估计式用到两个近似解 $\tilde{x}_n^*$ 和 $\tilde{x}_{2^k}^*$ 的差上, 我们就得到

$$\|\tilde{x}_{2^k}^*-\tilde{x}_n^*\|_C\leqslant M2^{k/2}\|\tilde{x}_{2^k}^*-\tilde{x}_n^*\|_{L^2},$$

其中 $2^{k-1}\leqslant n\leqslant 2^k$. 其次, 利用估计式(9)得到

$$\|\tilde{x}_{2^k}^*-\tilde{x}_n^*\|_{L^2}\leqslant\|\tilde{x}_{2^k}^*-x^*\|_{L^2}+\|x^*-\tilde{x}_n^*\|_{L^2}$$

$$\leqslant\frac{C}{2^{\frac{k}{2}+\alpha k}}+\frac{C}{n^{\frac{1}{2}+\alpha}}\leqslant\frac{C_2}{2^{\frac{k}{2}+\alpha k}},$$

因而

$$\|\tilde{x}_{2^k}^*-\tilde{x}_n^*\|_C\leqslant\frac{C_2}{2^{\alpha k}}. \tag{10}$$

特别, 当 $n=2^{k-1}$ 时就有

$$\|\tilde{x}_{2^k}^*-\tilde{x}_{2^{k-1}}^*\|_C\leqslant\frac{C_2}{2^{\alpha k}}.$$

由此可知, 级数

$$\tilde{x}_1^*(t)+\sum_{k=1}^{\infty}[\tilde{x}_2^{*k}(t)-\tilde{x}_2^{*k-1}(t)]=\lim_{k\to\infty}\tilde{x}_2^{*k}(t)=x^*(t)$$

一致收敛. 由(10)可知, 极限

$$\lim_{n\to\infty}\tilde{x}_n^*(t)=x^*(t)$$

一致存在. 而且,显然有

$$\|x^*-x_n^*\|_{\mathbf{C}}=O(n^{-\alpha}).$$

如果将积分方程(1)看作空间 $\mathbf{C}[0,1]$中的方程, 则应用定理1.3可得出比刚才指出的更强的结果,即如果核函数 $h(s,t)$(关于 s)和右端 $y(s)$ 满足指数为 $\beta>0$ 的 Lipschitz 条件, 则近似解一致收敛于精确解,并且

$$\|x^*-x_n^*\|_{\mathbf{C}}=O(n^{-\beta}\ln n).$$

4.4. 现在来看看核置换法. 所谓核置换法也就是用近似核,特别是用退化核去代替原来的核函数的一类解法. 按照这种方法,将积分方程(1)换成

$$\tilde{x}(s)-\lambda\int_0^1\tilde{h}(s,t)\tilde{x}(t)dt=y(s), \tag{11}$$

其中新的核 $\tilde{h}(s;t)$是核 $h(s,t)$的某种近似,而方程(11)的解是已知的.

将方程(11)写成与方程(1)在同一空间中的泛函方程

$$\widetilde{K}\tilde{x}\equiv\tilde{x}-\lambda\widetilde{H}\tilde{x}=y,$$

我们即把问题归结为在1.8最后所述的情形. 正如在那里所指出的,条件I归结为不等式

$$\|H-\widetilde{H}\|\leqslant\eta, \tag{12}$$

而条件II和III在 $\eta_1=\eta_2=0$ 时成立.

取有界可测函数空间 $\mathbf{L}^\infty(0,1)$作为空间 $\mathbf{X}=\widetilde{\mathbf{X}}$, 将(12)式改写为

$$\int_0^1 |h(s,t)-\tilde{h}(s,t)|\,dt \leqslant \eta \qquad (0\leqslant s\leqslant 1).$$

如果已知近似方程(11)的豫解式 $\tilde{\Gamma}(\lambda; s, t)$，则不难估计算子 $\tilde{K}^{-1}$ 的范数．因逆算子 $\tilde{K}^{-1}$ 具有形式

$$x=\tilde{K}^{-1}y,$$

$$z(s)=y(s)+\lambda\int_0^1 \tilde{\Gamma}(\lambda; s, t)y(t)\,dt,$$

则

$$\|\tilde{K}^{-1}\|\leqslant 1+|\lambda|B,$$

其中 B 由不等式

$$\int_0^1 |\tilde{\Gamma}(\lambda; s, t)|\,dt\leqslant B \qquad (0\leqslant s\leqslant 1)$$

确定．

利用 §1 中的估计式(19)，我们得到近似解接近于精确解的如下估计

$$|x^*(t)-\tilde{x}^*(t)|\leqslant|\lambda|\eta(1+|\lambda|B)\|x^*\|. \tag{13}$$

如果改变两个方程的作用，则得类似的估计．

由(13)推知，当核函数 $\tilde{h}(s,t)$ 一致收敛于 $h(s,t)$ 时，近似解一致收敛于精确解．

关于算子方程的一般理论，还可应用到求形如(7)式的解上，而且要求等式的左右两端仅仅在已知的点上相等．在这里我们不再叙述和研究这种方法了(见 Канторович[9])，因为在下一节对微分方程还有类似的方法要得到讨论．

§5. 对常微分方程的应用

5.1. 考察如下的边值问题

$$\frac{d^{2m}x}{dt^{2m}}-\lambda\left[p_1\frac{d^{2m-1}x}{dt^{2m-1}}+\cdots+p_{2m}x\right]=y \tag{1}$$

$$\begin{aligned} x(a)=x'(a)=\cdots=x^{(m-1)}(a)=0\\ x(b)=x'(b)=\cdots=x^{(m-1)}(b)=0 \end{aligned} \tag{2}$$

的所谓插值法(或叫重迭法)*).

我们来求形如

$$\tilde{x}(t)=(t-a)^m(b-t)^m\sum_{k=1}^{n}c_k t^{k-1} \tag{3}$$

的近似解. 易见, 对 $\tilde{x}(t)$ 来说边界条件(2)是满足的. 系数 c_1, $c_2, \cdots, c_n$ 由下述关系式确定:

$$\left\{\frac{d^{2m}\tilde{x}}{dt^{2m}}-\lambda\left[p_1\frac{d^{2m-1}\tilde{x}}{dt^{2m-1}}+\cdots+p_{2m}\tilde{x}\right]\right\}_{t=t_j}=y(t_j)$$

$$(j=1,2,\cdots,n). \tag{4}$$

也就是说使得(1)在一组已知结点 $t_1, t_2, \cdots, t_n$ 上成立而来决定诸系数 $c_1, c_2, \cdots, c_n$.

为了应用在§2 中所述的一般理论, 我们将微分方程(1)看成是空间 $\mathbf{X}=\mathbf{C}^{(2m)}[a,b]$中的泛函方程, 其中 $\mathbf{C}^{(2m)}[a,b]$表示由满足条件(2)的 $(2m)$次连续可微函数组成的空间. 该空间中算子的范数将在下面给出.

我们将形如(3)的函数的集合取为 $\tilde{\mathbf{X}}$.

其次, 用在区间$[a,b]$上的连续函数的集合 $\mathbf{C}[a,b]$作为空间 $\mathbf{Y}$, $\mathbf{C}[a,b]$中的范数按通常的方式定义.

最后, 用所有$(n-1)$次多项式构成的集合作为空间 $\tilde{\mathbf{Y}}$. 我们这样来定义从$\mathbf{Y}$到 $\tilde{\mathbf{Y}}$ 的映射 Φ, 对每个 $y\in\mathbf{Y}$, 命 $\tilde{y}=\Phi(y)$, 其中 $\tilde{y}\in\tilde{\mathbf{Y}}$是在结点 $t_1, t_2, \cdots, t_n$ 处取和 y 在这些点上相同的值的$(n-1)$次插值多项式. 正如大家所知道的, 当 $t_1, t_2, \cdots, t_n$ 是 Чебыщев 结点时,

$$\|\Phi\|\leqslant A\ln n+B. \tag{5}$$

---

*) 第一个插值法是由 Канторович[1]给出的. 基于 §§1—2 的定理的收敛性证明是在 Карпиловская[1]中得到的.

在 Gauss 结点的情形,

$$\|\Phi\| \leqslant A\sqrt{n}, \tag{6}$$

而如果结点是具有有界权 $\omega(t) \geqslant c > 0$ 的 n 阶正交多项式的根, 则

$$\|\Phi\| \leqslant An^{*)}. \tag{7}$$

所考察的边值问题等价于泛函方程

$$K_1 x \equiv Gx - \lambda Tx = y, \tag{8}$$

其中

$$Gx = \frac{d^{2m}x}{dt^{2m}}; \quad z = Tx, z(t) = \sum_{s=1}^{2m} p_s(t) x^{(2m-s)}(t). \tag{9}$$

算子 G 的逆算子是一个积分算子, 这个积分算子的核函数是微分算子$\frac{d^{2m}x}{dt^{2m}}$在条件(2) 下的 Green 函数.

Green 函数的存在可由微分方程

$$\frac{d^{2m}x}{dt^{2m}} = 0 \tag{10}$$

在条件(2)下具有唯一的零解这一事实推知. 那么,何以知道方程(10)在条件(2)下具有唯一零解呢?这是因为方程(10)的通解是任意的 $(2m-1)$ 次多项式, 而满足 (2) 的非零多项式应含有因子 $(t-a)^m(b-t)^m$, 这个因子的幂次为 $2m$.

因此,

$$x = G^{-1}y, x(s) = \int_a^b g(s, t) y(t) dt, \quad (x \in \mathbf{X}, y \in \mathbf{Y}).$$

函数 $g(s, t)$ 具有直到 $(2m-2)$ 阶的连续导数, 而它的 $2m-1$ 阶导数在 $s=t$ 处有跳跃, 由此

---

*) 关于(5)和(7), 见 Натансон-I 第 539 页和第544 页, 而估计式(6) 见于 Szegö 的书.

$$x^{(k)}(s)=\int_a^b \frac{\partial^k}{\partial s^k} g(s,t) y(t)\,dt \qquad (k=0,1,\cdots,2m-1),$$

于是

$$\max_s |x^{(k)}(s)| \leqslant A_k \|y\|_{\mathrm{C}} \quad (k=0,1,\cdots,2m-1). \qquad (11)$$

如果取 $A_{2m}=1$，则(11)式对 $k=2m$ 也成立.

要让 **X** 中的范数和 **Y** 中的范数相容，即使得算子 G 实现了这两个空间之间的等距，则应令

$$\|x\|_{\mathrm{X}}=\|Gx\|_{\mathrm{Y}}=\max_t |x^{(2m)}(t)|.$$

在这种情况下，(11)式可用于通过元素 x 在空间 **X** 中的范数来估计 x 的任意次导数:

$$\max_t |x^{(k)}(t)| \leqslant A_k \|x\| \qquad (k=0,1,2,\cdots,2m). \qquad (12)$$

我们指出，算子 G 乃是将空间 $\tilde{\mathbf{X}}$ 中的元素变成空间 $\tilde{\mathbf{Y}}$ 中的元素，即变成次数不超过$(n-1)$的多项式.

现在来估计作为从空间 $\mathbf{C}^{(2m)}[a,b]$到空间 $\mathbf{C}[a,b]$内的算子 T 的范数. 借助于(12)，可得

$$\|Tx\|=\max_t \left| \sum_{s=1}^{2m} p_s(t) x^{(2m-s)}(t) \right| \leqslant \left[\sum_{s=1}^{2m} B_s A_{2m-s} \right] \|x\|,$$

其中

$$B_s=\max_t |p_s(t)| \qquad (s=1,2,\cdots,2m).$$

这样一来，就有

$$\|T\| \leqslant \sum_{s=1}^{2m} B_s A_{2m-s}. \qquad (13)$$

如果已知方程(1)在条件(2)之下的 Green 函数，则用类似的

方法可以估计算子 K_1^{-1}的范数.

因为近似方程组(4)是用形如(3)的式子 $\tilde{x}$ 代入(1)并令两端在结点 $t_1, t_2, \cdots, t_n$ 处相等而得到的，这就等价于在这些结点处按已知值构造插值多项式方程，因此可将(4)写成

$$\tilde{K}_1\tilde{x} \equiv G\tilde{x} - \lambda\Phi T\tilde{x} = \Phi y. \tag{14}$$

将方程(14)与 2.3 中的方程(19)比较，可看出它是2.3中(19)的特例，于是条件 Ib 在 $\mu=0$ 时成立.

为了验证条件 IIb，即建立用 $\tilde{\mathbf{Y}}$中的元素逼近形如 Tx 的元素的可能性，我们来估计$\frac{d}{dt}Tx$. 由(12)，有

$$\left\|\frac{d}{dt}Tx\right\|_C = \max_t\left|\frac{d}{dt}\sum_{s=1}^{2m}p_s(t)x^{(2m-s)}(t)\right|$$

$$= \max_t\left|\sum_{s=0}^{2m}(p_s'(t)+p_{s+1}(t))x^{(2m-s)}(t)\right| \leqslant M\|x\|$$

$$(p_0(t)=p_{2m+1}(t)=0).$$

据 Jackson 定理，存在多项式 $\tilde{y}\in\tilde{\mathbf{Y}}$ 使得

$$\|Tx-\tilde{y}\| \leqslant \frac{AM\|x\|}{n}.$$

因此，条件 IIb 在 $\mu_1=O\left(\frac{1}{n}\right)$时成立.

至于条件 IIIb，如果方程(1)的右端函数 y 具有连续的导数，仍由 Jackson 定理可知条件 IIIb 在$\mu_2=O\left(\frac{1}{n}\right)$时成立.

还要指出，如果在方程(1)中 $p_1=0$(这一点经过适当的变量代换总可以达到)且系数 p_s 二次连续可微，则关于$\frac{d^2}{dt^2}Tx$ 进行前述的讨论，可得到 $\mu_1=O\left(\frac{1}{n^2}\right)$. 如果 y 也是二次连续可微函数，则也

有 $\mu_2=O\left(\frac{1}{n^2}\right)$.

利用§2 中的结果,考虑到此时对 μ_1, μ_2 及$\|\Phi\|$的估计,则由定理 2.1b 可知当 n 充分大时方程组(4)可解.

当结点是 Чебышев 结点或 Gauss 结点时,由定理 2.3b 可知近似解收敛于精确解,而且它们的估计分别如下:

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{\ln n}{n}\right),$$

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{\sqrt{n}}\right).$$

在上述情形,当 $p_1=0$ 时我们得到方程组(4)的可解性,如果用关于权 $\omega(t)\geqslant c>0$ 正交的 n 阶正交多项式的根作为结点,则近似解的误差估计式为

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{n}\right).$$

系数越光滑则收敛越快. 假定 $p_1, p_2, \cdots, p_{2m}$ 和 y 是 r 次可微的且它们的 r 阶导数满足指数为 α 的 Lipschitz 条件,利用定理 1.3 的注可知收敛速度由下式

$$\|x^*-\tilde{x}_n^*\|=O(\varepsilon\|P\|),\quad (P=G^{-1}\Phi G;\ \|P\|=\|\Phi\|) \tag{15}$$

确定,其中 ε 表征由 $\tilde{x}\in\tilde{\mathbf{X}}$去逼近解 x^* 的可能性. 因为

$$\|x^*-\tilde{x}\|_{\mathbf{X}}=\|G(x^*-\tilde{x})\|_{\mathbf{Y}}=\left\|\frac{d^{2m}x^*}{dt^{2m}}-\tilde{y}\right\|_{\mathbf{Y}}$$

$$(\tilde{y}=\tilde{G}x\in\tilde{\mathbf{Y}}),$$

故 ε 由用$(n-1)$次多项式去逼近解的 $2m$ 阶导数的阶来决定. 但不难用关于 r 的归纳法来证明:函数$\frac{d^{2m}x^*}{dt^{2m}}$具有满足指数为$\alpha$的 Lipschitz 条件的 r 阶导数. 因此,由 Jackson 定理,存在多项式 $\tilde{y}\in\tilde{\mathbf{Y}}$ 使得

$$\left\|\frac{d^{2m}x^*}{dt^{2m}}-\tilde{y}\right\|\leqslant\frac{D}{n^{r+\alpha}}.$$

因此，ε 具有阶$n^{-r-\alpha}$，而由(15)式，对 Чебышев 结点，Gauss 结点和关于下有界的函数为权的正交多项式的根为结点的情形，分别有如下估计式

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{\ln n}{n^{r+\alpha}}\right),\ \|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{n^{r+\alpha-1/2}}\right),$$

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{n^{r+\alpha-1}}\right).$$

注 1. 不必对证明的过程作实质性的改动，可以知道这个方法不仅对边值问题而且对初值问题（Cauchy 问题）也适用. 如果讨论的不是一个微分方程而是微分方程组，则讨论的方法也一样.

注 2. 这个近似方法的收敛性定理可看作是某个插值过程的收敛性定理. 这也就是说，多项式 $\tilde{x}_n^*(t)$可看作是 $x^*(t)$的插值多项式，后者可按边界条件及微分算子 Kx^* 在结点 $t_1, t_2, \cdots, t_n$ 处的值来构造.

注 3. 指出这一点是有趣的：对于等距结点，上述插值过程即使对于最简单的方程$\frac{d^2x}{dt^2}=y$ 也不收敛. 为了验证这一点，只要注意到下述事实就足够了，即普通的插值过程对等距结点也可能发散(见 Натансон-I 第 519 页).

5.2. 现在应用一般理论来研究 Галеркин 法和矩量法的收敛性问题*).

设要考察的问题是由微分方程 (1) 和在区间的端点上的齐次

*) 在 М. В. Келдыш[1] 的著作中第一次建立了 Галеркин 法在一般情形下的收敛性. 在 Н. М. врылов 和 Н. Н. Боголюбов 的早期著作中讨论了 Галеркин 法的特殊情形 Ritz 法及矩量法(见 Крылов 及 Михлин-II).

边界条件

$$M(x)=0 \tag{16}$$

所构成的边值问题. 矩量法乃是取满足条件(16)的函数系$\{\omega_k\}$而求形如

$$\tilde{x}(t)=\sum_{k=1}^{n} c_k\omega_k(t) \tag{17}$$

的近似解,并且 c_k 是用方程组

$$\int_a^b L\left(\sum_{k=1}^{n} c_k\omega_k\right)\zeta_j(t)\,dt=\int_a^b y(t)\zeta_j(t)\,dt \quad (j=1,2,\cdots,n)$$

来确定,其中 $L(x)$表示如下的微分表达式

$$L(x)=\frac{d^{2m}x}{dt^{2m}}-\lambda\sum_{s=1}^{2m} p_s\frac{d^{2m-s}x}{dt^{2m-s}}.$$

如果这时 $\zeta_j=\omega_j$, 则得 Галеркин 法.

5.3. 特别,考察方程

$$x''(t)-\lambda[p_1(t)x'(t)+p_2(t)x(t)]=y(t) \quad (|t|\leqslant 1) \tag{18}$$

在条件

$$x(-1)=x(1)=0 \tag{19}$$

下的解. 这时,作为 ω_k 可取

$$\omega_k(t)=t^{k-1}(1-t^2) \quad (k=1,2,\cdots).$$

方程(18)可写成

$$Gx-\lambda Tx=y, \tag{20}$$

其中 $Gx=x''$. 作为 $\mathbf{X}$ 可取由满足(19)的二次连续可微函数的(不完备的)酉空间,该空间的范数和内积分别是

$$\|x\|_{\mathbf{X}}=\left[\int_{-1}^{1}|x''(t)|^2\,dt\right]^{1/2},$$

$$(x_1,x_2)_{\mathbf{X}}=\int_{-1}^{1}x_1''(t)x_2''(t)\,dt.$$

取由所有的连续函数构成的集合为 $\mathbf{Y}$, 而 $\mathbf{Y}$ 中的距离是借助于算子 G 按空间 $\mathbf{X}$ 中的距离引入的, 即

$$\|y\|_{\mathbf{Y}}=\|G^{-1}y\|_{\mathbf{X}}=\left[\int_{-1}^{1}|y(t)|^{2}dt\right]^{1/2},$$

$$(y_1, y_2)_{\mathbf{Y}}=\int_{-1}^{1}y_1(t)y_2(t)dt.$$

取形如(17)的函数空间的子空间为 $\tilde{\mathbf{X}}$, 而 $\tilde{\mathbf{Y}}$ 取为 $G(\tilde{\mathbf{X}})$. 易见, $\tilde{\mathbf{Y}}$ 乃由所有 $(n-1)$ 次的多项式全体组成. 最后, 命 Φ 是由 $\mathbf{Y}$ 到 $\tilde{\mathbf{Y}}$ 上的正交投影算子. 这时, 如果 $\zeta_1, \zeta_2, \cdots, \zeta_n$ 是 $\tilde{\mathbf{Y}}$ 中的线性独立的元素组, 则等式 $\Phi y=\mathbf{0}$ 等价于等式组 $(y, \zeta_j)=0(j=1, 2, \cdots, n)$.

在所作的假定下, 矩量法的方程等价于泛函方程

$$G\tilde{x}-\lambda\Phi T\tilde{x}=\Phi y. \tag{21}$$

也就是说, 近似方程是由和 2.3 中相应的特殊方法构成的, 因而条件 Ib 在 $\mu=0$ 时成立.

为了验证条件 IIb, 应证明可由 $\tilde{\mathbf{Y}}$ 中的元素去逼近元素 Tx. 在此先指出一点, 即对 $\mathbf{X}$ 中的函数可由它的范数来估计这个函数本身和它的一阶导数的极大值:

$$\begin{aligned}|x'(t)|&=\frac{1}{2}\left|\int_{-1}^{1}[x'(t)-x'(\tau)]d\tau\right|\\&=\frac{1}{2}\left|\int_{-1}^{1}\left[\int_{\tau}^{t}x''(u)du\right]d\tau\right|\\&\leqslant\frac{1}{2}\int_{-1}^{1}\left[\int_{\tau}^{t}|x''(u)|^{2}du\right]^{1/2}\left[\int_{\tau}^{t}du\right]^{1/2}dt\leqslant A\|x\|,\end{aligned}$$

$$|x(t)|\leqslant\int_{-1}^{t}|x'(\tau)|d\tau\leqslant B\|x\|.$$

这时, 如果(18)的系数是连续可微的, 则有

$$\left\|\frac{d}{dt}Tx\right\|_{\mathbf{Y}}=\|p_1x''+(p_1+p_2')x'+p_2x\|_{\mathbf{Y}}\leqslant C\|x\|.$$

由此易知函数 Tx 满足 指数为$\frac{1}{2}$的 Lipschitz 条件

$$|(Tx)(t_1)-(Tx)(t_2)|=\left|\int_{t_1}^{t_2}\frac{d}{dt}Txdt\right|$$

$$\leqslant\left\|\frac{d}{dt}Tx\right\|_{\mathbf{Y}}|t_2-t_1|^{1/2},$$

故可用多项式 $\tilde{y}\in\tilde{\mathbf{Y}}$ 去逼近 Tx, 使得

$$\|Tx-\tilde{y}\|\leqslant\frac{D\|x\|}{\sqrt{n}}.$$

因此，可以认为 $\mu_1=O\left(\frac{1}{\sqrt{n}}\right)$.

假定右端的导数是平方可积的，由类似的论证可得知

$$\mu_2=O\left(\frac{1}{\sqrt{n}}\right).$$

如果 $p_1=0$, 则用同样的方法可以估计$\frac{d^2}{dt^2}Tx$ 并得到

$$\mu_1=O\left(\frac{1}{n\sqrt{n}}\right).$$

应用定理 2.3b, 可得到近似解序列收敛于精确解的速度的阶

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{\sqrt{n}}\right).$$

如果 $p_1=0$ 而 y'' 是平方可积的，则

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{n\sqrt{n}}\right).$$

如果 p_1, p_2 和 y 具有更高阶的导数，则近似解误差下降的阶就更高，因为这时 x^* 多次可微，故可用多项式更精确地去逼近它.

在空间 $\mathbf{X}$ 中近似解总是收敛于精确解的，也就是说，近似解本身和它的一阶导数分别一致地收敛于精确解和精确解的一阶导数，近似解的二阶导数平均收敛于精确解的二阶导数.

5.4. 我们仍来考察方程(18),不过这时空间$\mathbf{X}$和$\mathbf{Y}$中的范数是这样定义的:

$$\|x\|_{\mathbf{X}}=\max_{|t|\leqslant 1}|x''(t)|,\ \|y\|_{\mathbf{Y}}=\max_{|t|\leqslant 1}|y(t)|,$$
$$(x\in\mathbf{X},\ y\in\mathbf{Y}).$$

显然,这样定义 的范数是相容的,于是$\|G\|=\|G^{-1}\|=1$. 算子Φ仍象以前那样定义.因而可将近似方程看成是以前建立方程的特例,因此 $\mu=0$. 现在我们来估计Φ的范数. 设 $\zeta_1,\zeta_2,\cdots,\zeta_n$ 是以1为权函数的正交多项式(Legendre 多项式). 由$\mathbf{Y}$到$\tilde{\mathbf{Y}}$上的投影算子Φ这时可写成

$$\Phi y=\sum_{k=1}^{n}(y,\zeta_k)\zeta_k.$$

但根据连续函数按 Legendre 多项式展开的众所周知的估计(例如,可见 Агаханов 和 Натансон[1])

$$\max_{|t|\leqslant 1}\left|\sum_{k=1}^{n}(y,\zeta_k)\zeta_k(t)\right|\leqslant A\sqrt{n},$$

于是

$$\|\Phi\|\leqslant A\sqrt{n}. \tag{22}$$

为了用多项式去逼近 Tx, 也就是用$\tilde{\mathbf{Y}}$中的元素去逼近 Tx, 我们利用不等式

$$\left\|\frac{d}{dt}Tx\right\|_{\mathbf{Y}}\leqslant C\|x\|.$$

所以可求得 $\tilde{y}\in\tilde{\mathbf{Y}}$ 使得

$$\|Tx-\tilde{y}\|\leqslant\frac{D\|x\|}{n}.$$

因而 $\mu_1=O\left(\frac{1}{n}\right)$$\left(\text{如果 } p_1=\mathbf{0}\text{, 则可取 } \mu_1=O\left(\frac{1}{n^2}\right)\right)$. 最后,

$$\mu_2=\frac{E_n(y)}{\|y\|}.$$

根据定理 2.2b, 由(22)式可得

$$\|x^*-\tilde{x}_n^*\|=(\mu_1+\mu_2)O(\sqrt{n}).$$

如果 $E_n(y)\sqrt{n}\to 0$, 则近似解及其一阶和二阶导数分别一致收敛于精确解及其一阶和二阶导数. 此外, 如果 $p_1=0$ 而 y 是二次可微的, 则

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{n\sqrt{n}}\right).$$

5.5. 作为更进一步的应用, 我们考察方程

$$L(x)\equiv\frac{d}{dt}\left[p\frac{dx}{dt}\right]-\lambda\left\{\frac{d}{dt}[qx]+rx\right\}=y \tag{23}$$

在条件

$$x(0)=x(1)=0 \tag{24}$$

下的解. 我们假定方程 (23) 中的函数 p 和 q 连续可微而且 $p(t)>0$. 此外, 还认为 λ 使得上述边值问题具有唯一的解.

按 Галеркин 法, 我们来求形如

$$\tilde{x}(t)=\sum_{k=1}^{n}c_k\omega_k(t)\qquad(0\leqslant t\leqslant 1) \tag{25}$$

的近似解, 其中 $\omega_k(k=1,2,\cdots)$ 是满足条件(24)的连续可微函数:

$$\omega_k(0)=\omega_k(1)=0\qquad(k=1,2,\cdots).$$

于是, ω_k的导数和恒为 1 的函数一起构成完全函数系. 设这个完全系$\{\omega_k'\}$以 p 为权而正交:

$$\int_0^1 p(t)\omega_j'(t)\omega_k'(t)dt=\begin{cases}0, & j\neq k,\\ 1, & j=k\end{cases}\ (j,k=1,2,\cdots). \tag{26}$$

如果将函数 $\omega_0'(t)=\frac{1}{p(t)}\left[\int_0^1\frac{ds}{p(s)}\right]^{-1/2}$补充到$\{\omega_k'\}$中去, 则(26)式对 $k=0, j=0$ 也成立, 于是函数系$\{\omega_k'\}$ $(k=0,1,2,\cdots)$构成完全

正交函数系.

设$\mathbf{X}$是由满足(24)的二次连续可微函数组成的空间，则方程(23)和条件(24)合在一起可看成是空间$\mathbf{X}$中的一个泛函方程

$$Gx-\lambda Tx=y, \tag{27}$$

其中,空间$\mathbf{X}$中的范数是这样来定义的:

$$\|x\|=\left[\int_0^1 p(t)|x'(t)|^2dt\right]^{\frac{1}{2}}; \quad (x_1, x_2)=\int_0^1 p(t)x_1'(t)x_2'(t)dt.$$

空间$\mathbf{Y}$由连续函数组成，$\mathbf{Y}$中的距离按它和空间$\mathbf{X}$中距离的相容性条件引入:

$$\|y\|_{\mathbf{Y}}=\|G^{-1}y\|_{\mathbf{X}}{}^{*)}, \quad (y_1, y_2)_{\mathbf{Y}}=(G^{-1}y_1, G^{-1}y_2)_{\mathbf{X}}.$$

$\tilde{\mathbf{X}}$取为所有形如(25)的元素的集合，而$\tilde{\mathbf{Y}}$取为所有形如

$$y=\sum_{k=1}^{n}c_kG\omega_k \tag{28}$$

的元素的集合. 最后,Φ定义为从$\mathbf{Y}$到$\tilde{\mathbf{Y}}$上的正交投影算子，从而$\|\Phi\|=1$. 等式$\Phi y=\mathbf{0}$等价于$(y, G\omega_k)=0(k=1, 2, \cdots, n)$，而由$\mathbf{Y}$中内积的定义,后一组等式又等价于

$$\begin{aligned}(G^{-1}y,\ \omega_k)_{\mathbf{X}}&=\int_0^1 p(t)\frac{d}{dt}(G^{-1}y)\omega_k'(t)dt\\&=-\int_0^1 y(t)\omega_k(t)dt=0,\\&(k=1, 2, \cdots, n).\end{aligned}$$

因此,Галеркин 法的方程

$$\int_0^1[L(x)(t)-y(t)]\omega_j(t)dt=0, \quad (j=1, 2, \cdots, n)$$

不是别的,正是表示式

---

*) 元素$x=G^{-1}y$是方程$\frac{d}{dt}\left(p\frac{dx}{dt}\right)=y$的满足条件(24)的解.

$$\Phi[Gx-\lambda Tx]=\Phi y.$$

也就是说，是 2.3 中所建立的近似方程的特殊形式.

为了应用 §2 中的结果，注意这时有 $\mu=0$. 我们再来估计 μ_1. 为此，利用 2.1 中的注(不等式(6))，亦即我们将用$\tilde{\mathbf{X}}$中的元素在空间$\mathbf{X}$中的距离意义下去逼近元素 $z=G^{-1}Tx(\|x\|\leqslant 1)$. 因为 $Gz=Tx$，即

$$\frac{d}{dt}\left[p\frac{dz}{dt}\right]=\frac{d}{dt}[qx]+rx,$$

从 0 到 $u\leqslant 1$ 积分这个等式，得

$$\begin{aligned}p(u)z'(u)&=q(u)x(u)+\int_0^u r(t)x(t)dt+C\\&=q(u)\int_0^u x'(t)dt+\int_0^u\left(\int_t^u r(\tau)d\tau\right)x'(t)dt+C'\\&=\int_0^1 h(u,t)x'(t)dt+C,\end{aligned}$$

其中
$$h(u,t)=\begin{cases}q(u)+\int_t^u r(\tau)d\tau & (t\leqslant u),\\ 0 & (t>u).\end{cases}$$

为了消去常数 C，在所得等式两端先除以 $p(u)$ 然后再积分，则有

$$0=\int_0^1\left[\int_0^1\frac{1}{p(u)}h(t,u)du\right]x'(t)dt+C\int_0^1\frac{du}{p(u)}.$$

由此定出 C，再代入前面的等式得到

$$z'(u)=\int_0^1 K(u,t)x'(t)dt,$$

其中

$$K(u,t)=\frac{1}{p(u)}h(u,t)-\frac{1}{p(u)\int_0^1\frac{ds}{p(s)}}\int_0^1\frac{1}{p(s)}h(s,t)ds.$$

我们指出，显然有

$$\int_0^1 K(u,t)du=0\qquad(0\leqslant t\leqslant 1).$$

可用关于$\{\omega_j'(u)\}$ $(j=1, 2, \cdots)$的Fourier级数的部分和

$$K_n(u, t)=\sum_{j=1}^{n}\alpha_j(t)\omega_j'(u)$$

去逼近核函数 $K(u, t)$，因为核函数正交于函数$\frac{1}{p(u)}$而函数系$\{\omega_j'\}$ $(j=0, 1, 2, \cdots)$是完全的，其中

$$\alpha_j(t)=\int_0^1 p(u)K(u, t)\omega_j'(u)du.$$

这样一来，当 n 增大时，单调地有

$$\int_0^1 p(u)[K(u, t)-K_n(u, t)]^2du \xrightarrow[(n\to\infty)]{} 0 \quad (0\leqslant t\leqslant 1). \qquad (29)$$

由等式

$$\tilde{x}'(u)=\int_0^1 K_n(u, t)x'(t)dt=\sum_{j=1}^{n}c_j\omega_j'(u)$$

$$\left(c_j=\int_0^1 \alpha_j(t)x'(t)dt\right),$$

确定 $\tilde{x}$，我们就得到所需要的近似. 事实上，

$$\|z-\tilde{x}\|^2=\int_0^1 p(u)|z'(u)-\tilde{x}'(u)|^2du$$

$$=\int_0^1 p(u)\left|\int_0^1 [K(u, t)-K_n(u, t)]x'(t)dt\right|^2du$$

$$\leqslant\int_0^1 p(t)|x'(t)|^2dt\int_0^1\left[\int_0^1 |K(u, t)-K_n(u, t)|^2p(u)du\right]\frac{dt}{p(t)}$$

$$=\eta_n^2\|x\|^2,$$

其中

$$\eta_n^2=\int_0^1\frac{dt}{p(t)}\int_0^1 [K(u, t)-K_n(u, t)]^2p(u)du \xrightarrow[(n\to\infty)]{} 0.$$

用类似的方法可得知 μ_2 充分小. 应该证明可用 $\tilde{\mathbf{X}}$ 中的元素

去逼近元素 $z=G^{-1}y$. 由于

$$\frac{d}{dt}\left(p\frac{dz}{dt}\right)=y, \qquad p\frac{dz}{dt}=\int y\,dt=F,$$

所以选择 F 使得 $\int_0^1 \frac{F(t)}{p(t)}dt=0$, 这时 $\frac{dz}{dt}=\frac{F}{p}$ 可用形如 $\sum_{k=1}^{n} c_k\omega_k'$ 的和来逼近.

如果取三角函数系为 ω_k, 则还可以估计近似的阶. 亦即, 设

$$\omega_k(t)=\sin k\pi t \qquad (k=1, 2, \cdots).$$

核函数 $K(u, t)$ 这时应由形如

$$\sum_{k=1}^{n}\alpha_k(t)\cos k\pi u$$

的和去逼近, 其中

$$\alpha_k(t)=2\int_0^1 K(u, t)\cos k\pi u\,du$$

是囿变函数的 Fourier 系数, 具有阶 $1/k$, 而偏差

$$\int_0^1 [K(u, t)-K_n(u, t)]^2 du=2\sum_{k=n+1}^{\infty}|\alpha_k(t)|^2=O\left(\frac{1}{n}\right).$$

因此 $\mu_1=O\left(\frac{1}{\sqrt{n}}\right)$. 同样, 容易验证 $\mu_2=O\left(\frac{1}{\sqrt{n}}\right)$.

因此, 在所考察的情形, 由定理 2.3b 可得估计

$$\|x^*-\bar{x}_n^*\|=O\left(\frac{1}{\sqrt{n}}\right),$$

亦即近似解按空间 **X** 中的距离收敛于精确解, 同样还保证以此速度一致收敛于精确解.

如果微分表达式 $L(x)$ 是自共轭的, 即若 $q=0$, 则 Галеркин 法就是 Ritz 法. 此时, 假定 $p(u)K(u, t)$ 具有关于 u 的囿变的一阶导数(这只要函数 p 和 r 具有类似的性质就可以), 则 Fourier

系数的阶为 $1/k^2$，故得到估计

$$\mu_1 = O\left(\frac{1}{n\sqrt{n}}\right)$$

以及近似解收敛于精确解的速度的相应估计.

当解可用函数系很好地逼近时，利用定理 1.3 的注，我们还可得到近似解的更高的收敛速度.

§ 6. 对椭圆型方程边值问题的应用

6.1. 这里，我们来考察用近似方法的一般理论于某些椭圆型方程边值问题的近似解法. 特别，考察方程

$$\Delta u + \lambda a u = v \tag{1}$$

在条件

$$u|_\Gamma = 0 \tag{2}$$

下的边值问题，其中(1)中的 u 属于由曲线 Γ 所围成的区域 D. 这时，除了近似方法一般理论中的假定外，我们还要用到数学物理方程和逼近论中的某些结果. 这些结果我们只引用而不再去证明了.

定理(Гюнтер-Корн). 如果 u 是方程

$$\Delta u = f, \ u|_\Gamma = 0 \tag{3}$$

的解，其中 Γ 是光滑曲线而 f 满足指数为 β 常数为 M 的 Lipschitz 条件($f \in \mathrm{Lip}_M \beta$)，则 $u(x, y)$ 具有二阶偏导数且这些偏导数具有以任何 $\beta' < \beta$ 为指数的 Lipschitz 条件，亦即有

$$\frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y^2} \in \mathrm{Lip}_{M'} \beta' \qquad (M' \leqslant CM, C = C(\beta, \beta'))^{*)}.$$

众所周知，对一个变量的函数而言，函数具有某种可微性质就同时保证了有用代数多项式或三角多项式来逼近该函数的可能

---

*) 见 Гюнтер[1].

性．此时，如果这个函数在区间端点 a, b 处变为零，则可以使得逼近函数在端点也变成零，特别，在用代数多项式去逼近时就可以要求这个多项式包含因子 $(t-a)(b-t)$.

对多变量函数，类似的定理也成立．下面我们给出其中两个定理，一个建立相应函数系的完全性，另一个给出逼近精确性的特征.

定理(完全性)*)．设 Γ 是由方程

$$\omega(x, y)=0$$

所确定的闭曲线，其中 ω 是分段连续可微函数而且

$$\omega(x, y)>0\ ((x, y)\in D),\ \mathrm{grad}\,\omega(x, y)\neq 0 \quad ((x, y)\in \Gamma), \tag{4}$$

这里的 D 表示由曲线 Γ 所界的区域．那么，形如

$$\tilde{u}(x, y)=\omega(x, y)P(x, y) \quad (P \text{ 为多项式})$$

的函数系在空间 $\mathring{W}_2^{(1)}$ 中是完全的．这里的 $\mathring{W}_2^{(1)}$ 表示由 $W_2^{(1)}$ 中在 Γ 上变为零的那些函数组成的空间.

定理(Харрик)**)．设函数 ω 满足完全性定理的条件，此外，它 k 次连续可微且 k 阶导数满足 Lipschitz 条件．如果函数 u 也是 k 次连续可微的且它的 k 阶导数满足指数为 α 的 Lipschitz 条件，则存在次数不超过 n 的多项式序列 $\{P_n\}$，使得

$$\|u-\omega P_n\|_{C^{(r)}}=O\left(\frac{1}{n^{k+\alpha-r}}\right), \tag{5}$$

其中

$$\|u\|_{C^{(r)}}=\sum_{i=0}^{r}\sum_{\nu_1+\nu_2=i}\max_D\left|\frac{\partial^i u}{\partial x^{\nu_1}\partial y^{\nu_2}}\right|. \tag{6}$$

6.2. 现在来看在这一节的开始所提出的边值问题．我们将寻求这个问题的形如

---

*) 见 Канторович 和 Крылов，第 296 页.

**) 见 Харрик [2].

$$\tilde{u}(x, y)=\omega(x, y)P(x, y)$$

的近似解，其中P是次数不超过n的多项式，而ω是满足 Харрик 定理当$k=1$时的条件.

多项式P的系数(共N个)由方程组

$$\iint_D (\Delta\tilde{u}+\lambda a\tilde{u})\zeta_k\,dx\,dy=\iint_D v\zeta_k\,dx\,dy \quad (k=1, 2, \cdots, N) \tag{8}$$

确定(关于ζ_k的选择将在下面说明).

我们引进空间$\mathbf{U}$，这是由在Γ上变为零的二次连续可微函数u组成的，$\mathbf{U}$中的距离和内积由公式

$$\|u\|=\left[\iint_D |\Delta u|^2 dx\,dy\right]^{1/2}, \quad (u_1, u_2)=\iint_D \Delta u_1\Delta u_2\,dx\,dy$$

定义. 取形如$\tilde{u}$的函数集合作为子空间$\tilde{\mathbf{U}}$.

取形如$v=\Delta u(u\in\mathbf{U})$的函数组成的空间作为$\mathbf{V}$，换言之，$\mathbf{V}$乃是由使得方程$\Delta u=v$具有属于$\mathbf{U}$的解的函数$v$所组成的空间. 空间$\mathbf{V}$中的距离取为 Hilbert 空间$\mathbf{L}^2(D)$中的距离. 最后，取所有形如 $\tilde{v}=\Delta\tilde{u}(\tilde{u}\in\tilde{\mathbf{U}})$的函数组成空间$\tilde{\mathbf{V}}$.

在这样的定义之下，已知的边值问题可看作形如

$$Gu+\lambda Tu=v \qquad (Gu=\Delta u; Tu=au) \tag{9}$$

的泛函方程. 这时，空间$\mathbf{U}$和$\mathbf{V}$中的距离这样来使其相容，即算子G实现从$\mathbf{U}$到$\mathbf{V}$的等距映射.

作为$\zeta_1, \zeta_2, \cdots, \zeta_N$可取$\tilde{\mathbf{V}}$中任意的线性独立函数组，例如可取

$$\zeta_{ij}=\Delta[\omega(x, y)x^i y^j] \qquad (i+j\leqslant n).$$

此时，如果将Φ看作是从$\mathbf{V}$到$\tilde{\mathbf{V}}$的正交投影变换，则$\Phi v=0$等价于

$$(v, \zeta_k)=\iint_D v\zeta_k\,dx\,dy=0 \qquad (k=1, 2, \cdots, n).$$

因此，方程组(8)在新的记号下可写成

$$\Phi G\tilde{u}-\Phi T\tilde{u}=\Phi v, \tag{10}$$

亦即近似方程构造为对应于 2.3 中的特殊形式. 因此，条件 Ib 在 $\mu=0$ 时成立.

为了验证 IIb，我们来证明，对于 $u\in\mathbf{U}$，有 $Tu=au\in\mathrm{Lip}\beta$，其中当 $m=2$ 时 $\beta<1$ 而当 $m=3$ 时 $\beta<1/2$.

事实上，因 $u\in\mathbf{U}$，故 $u\in\mathbf{W}_2^{(2)}$. 此时由嵌入定理的注 (XI4.4) 可知 $u\in\mathrm{Lip}\beta$，其中

$$\beta<1 \quad (\text{当 } m=2), \qquad \beta<1/2 \ (\text{当 } m=3).$$

由此，从 Гюнтер-Корн 定理推知 $z=G^{-1}Tu=\Delta^{-1}au$ 具有属于 $\mathrm{Lip}\,\beta'$ 的二次导数，其中 $\beta'<\beta$，即当 $m=2$ 时 $\beta'<1$，当 $m=3$ 时 $\beta'<1/2$，并且有

$$\left\|\frac{\partial^2 z}{\partial x^2}\right\|_{\mathrm{Lip}\beta'}\leqslant C\|u\|.$$

在这种情况下，按 Харрик 定理，函数 z 连同它的二阶导数可用形如 $\tilde{u}$ 的函数以 $1/n^{\beta'}$ 阶来逼近，而这就保证有估计式

$$\|G^{-1}Tu-\tilde{u}\|\leqslant\frac{C_1}{n^{\beta'}}.$$

因此，条件 IIb 在 $\mu_1=O(1/n^{\beta'})$ 下成立.

如果 v 属于 $\mathrm{Lip}\beta$，则也有类似的结果. 特别，如果 $v\in\mathbf{W}_2^{(2)}$，则进行类似的讨论可得 $\mu_2=O(1/n^{\beta'})$.

综上所述，可知这时以上述函数为基本函数的矩量法是收敛的，并且按空间 $\mathbf{U}$ 中的距离其收敛速度为

$$\|u^*-\tilde{u}_n^*\|_{\mathrm{U}}=O(1/n^{\beta'}).$$

由此易知 $\tilde{u}_n^*$ 是一致收敛于 u^* 的. 事实上，用按 Green 函数表示 u 的表达式得

$$\left|u^*(x,y)-\tilde{u}_n^*(x,y)\right|=\left|\iint\limits_D g(x,y)\Delta(u^*-\tilde{u}_n^*)\,dx\,dy\right|\leqslant$$

$$\leqslant\left[\iint_D |g(x,y)|^2\,dx\,dy\right]^{1/2}\|u^*-\tilde{u}_n^*\|=O(1/n^{\beta'}).$$

如果解具有更高阶的可微性质，则还可以得到更快的收敛速度.

6.3. 如果不考虑收敛速度的精确的阶，则还可以在更一般的条件下来建立 Галеркин 法的收敛性. 这些条件较之 6.2 中矩量法的条件更为一般.

考察方程

$$\Delta u+\lambda\left(au+b\frac{\partial u}{\partial x}+c\frac{\partial u}{\partial y}\right)=v \tag{11}$$

在条件

$$u|_\Gamma=0 \tag{12}$$

下的边值问题，假定其中的系数 a,b,c 是连续可微函数.

我们来求这个问题的如同在 6.2 中已指出的那种形式的解，其中的系数由 Галеркин 方程组

$$\iint_D\left[\Delta\tilde{u}+\lambda\left(a\tilde{u}+b\frac{\partial\tilde{u}}{\partial x}+c\frac{\partial\tilde{u}}{\partial y}\right)\right]\zeta_k\,dx\,dy=\iint_D v\zeta_k\,dx\,dy$$

$$(k=1,2,\cdots,N)$$

来确定. 在这里函数 ζ_k 取 $\zeta_k=\omega(x,y)P_k(x,y)$ 的形式，其中 P_k 是次数不超过 n 的多项式. 另外，这时我们用另一种方法引入空间 $\mathbf{U}$ 中的距离:

$$(u_1,\ u_2)=-\iint_D u_1\Delta u_2\,dx\,dy=\iint_D\left[\frac{\partial u_1}{\partial x}\frac{\partial u_2}{\partial x}+\frac{\partial u_1}{\partial y}\frac{\partial u_2}{\partial y}\right]dx\,dy,$$

$$\|u\|^2=-\iint_D u\Delta u\,dx\,dy=\iint_D\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]dx\,dy.$$

对空间 $\mathbf{V}$，也相应地引入距离:

$$(v_1,v_2)_{\mathbf{V}}=(G^{-1}v_1,G^{-1}v_2)_{\mathbf{U}},\ \|v\|_{\mathbf{V}}=\|G^{-1}v\|_{\mathbf{U}}.$$

即若 u_1 和 u_2 是方程 $\Delta u=v_1$ 和 $\Delta u=v_2$ 的属于 $\mathbf{U}$ 的解，则可取

$$(v_1, v_2) = -\iint_D u_1 v_2 dx dy = -\iint_D u_2 v_1 dx dy,$$

$$\|v_1\|^2 = -\iint_D u_1 v_1 dx dy.$$

如同前面的讨论一样，边值问题(11)—(12)可写成空间 $\mathbf{U}$ 中的一个泛函方程

$$Gu + \lambda Tu = v \qquad \left(Gu = \Delta u, Tu = au + b\frac{\partial u}{\partial x} + c\frac{\partial u}{\partial y}\right).$$

而其近似方程则具有如下形式：

$$G\tilde{u} + \lambda \Phi T \tilde{u} = \Phi v.$$

要利用定理 1.6，我们先来证明 $G^{-1}T$ 是紧的.

容易知道，$\mathbf{U}$ 和 $\mathring{\mathbf{W}}_2^{(1)}$ 中的一个稠密子集线性等距. 为了证实这一点，只要取由泛函 $\int_\Gamma u(x, y) d\Gamma$ 所定义的距离作为 $\mathbf{W}_2^{(1)}$ 中一个可能的距离即可，

$$\|u\|_{\mathbf{W}_2^{(1)}} = \left|\int_\Gamma u(x, y) d\Gamma\right| + \left\{\iint_D |\operatorname{grad} u|^2 dx dy\right\}^{1/2}$$

$$= \left\{\iint_D \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] dx dy\right\}^{1/2} = \|u\|_{\mathbf{U}},$$

因而，T 可看成从空间 $\mathring{\mathbf{W}}_2^{(1)}$ 到空间 $\mathbf{L}^2$ 的连续线性算子. 由 Бернштейн-Ладыженская 不等式，作为从 $\mathbf{L}^2$ 到 $\mathring{\mathbf{W}}_2^{(1)}$ 中算子的逆算子 Δ^{-1} 是连续的（见 Ладыженская, 第二章 §6）. 因此 $\Delta^{-1}T \in B(\mathring{\mathbf{W}}_2^{(1)}, \mathring{\mathbf{W}}_2^{(2)})$. 由 $\mathring{\mathbf{W}}_2^{(2)}$ 到 $\mathring{\mathbf{W}}_2^{(1)}$ 的嵌入算子是紧的(XI. 3.5)，因此，视为从 $\mathring{\mathbf{W}}_2^{(1)}$ 到 $\mathring{\mathbf{W}}_2^{(1)}$ 的算子 $\Delta^{-1}T$ 也是紧的.

由完全性定理，投影算子序列 $P_n = G^{-1}\Phi G$ 收敛于 $\mathbf{U}$ 上的单位算子. 因此，由定理 1.6 知 Галеркин 法在空间 $\mathbf{U}$ 中收敛. 由

此还得知近似解一致收敛于精确解.

以上的结论和证明不经任何本质上的更动和补充均可应用到具有更多变元的边值问题上去.

最后指出，本章所述方法的其他有趣的应用还有一系列的著作可资参考，特别是可参见 Каландия[1]，Владимиров[1]，Филиппов 和 Рябенький.

第十五章　最速下降法

最速下降法乃是目前解无条件极值问题的最普遍使用的方法之一. 这个方法最初是作为解线性泛函方程和求线性算子的特征值的变分法而被研究的. 如同所有的变分法一样，是将解方程(求特征值)问题化为求某个在整个空间上给定的特殊形式的泛函的极值问题. 但有比上面讲到的更为一般的求泛函极小值的合适的方法. 使用这种方法在经过不多的几步之后，解就被限于稳定点(或称驻点)的充分小的邻域之内(如果泛函是凸的，则限于极小点的邻域之中).

这种方法的要点在于：构造逼近极小元素的近似序列，而从一个近似到与之相邻的下一个近似乃是沿给定泛函的最速下降方向来进行的. 最速下降法也就故此而得名.

最速下降法的思想是属于 Cauchy 的，他当时对于有限维空间考察最速下降法. Канторович 对二次泛函研究了最速下降法的一般情形 (见 Канторович[7], [9]). 在下面(§1 和 §2)一般定理的证明中，利用了 М. Щ. Бирман 的著作[1]和[2]. 最速下降法的某些变形及其对种种问题的应用可参见 Вайнберг, Красносельский, Любич 和 Майстровский[1]. 最速下降法对椭圆型微分方程解的应用(§3)乃取材于 Л. Цах 的学位论文(1955 年). 在 §4 中，应用前面的著作于条件梯度法的研究. 这方面可参见 Демьянов 和 Рубинов 的著作.

§1. 线性方程的解

1.1. 设 Φ 是给定在赋范空间 **X**(实的或复的)上的实的非线性泛函. 还假定泛函 Φ 在 **X** 上是下有界的. 这时可提出这样的问

题：求元素 $x^*\in\mathbf{X}$（如果这个元素存在的话）使泛函达到极小，亦即使

$$\Phi(x)\geqslant\Phi(x^*)\qquad(x\in\mathbf{X}).$$

为了解决这个问题通常采用下面的办法：用某种方法构造出泛函 Φ 的"极小化"序列 $\{x_n\}$，即使得有

$$\lim_{n\to\infty}\Phi(x_n)=\inf_{x\in\mathbf{X}}\Phi(x).$$

在某些情况下上述的序列 $\{x_n\}$ 可实际构造出来而且收敛于某一元素 x^*. 如果泛函 Φ 是连续的，则 x^* 正是问题的解.

最速下降法乃是构造这样的序列 $\{x_n\}$ 的方法，即使得这种序列对于充分广的一类泛函是极小化序列. 这个方法可对沿方向可微的泛函 Φ 来描述. 我们现在来引入相应的定义.

设 x 是 $\mathbf{X}$ 中指定的元素. 考察从 x 出发沿方向 z 的射线，即考虑形如 $x+\alpha z$ 的所有元素，其中 α 是非负实数，而 z 是刻划射线方向的非零元素 $(x\in\mathbf{X})$*⁾. 泛函 Φ 局限在已给定的射线上可看成实变量的函数，若以 $\varphi(\alpha;x,z)$ 表示该函数，则按刚才所说，即有

$$\varphi(\alpha;x,z)=\Phi(x_0+\alpha z)\qquad(\alpha\geqslant0).$$

这时，我们自然就将

$$\frac{\partial\Phi}{\partial z}(x)=\frac{1}{\|z\|}\varphi'(\alpha;x,z)\Big|_{\alpha=0}=\frac{1}{\|z\|}\lim_{h\to0+}\frac{\Phi(x+hz)-\Phi(x)}{h}$$

称为泛函 Φ 在点 x 处沿方向 z 的导数. 我们假定导数 $\frac{\partial\Phi}{\partial z}(x)$ 对任意方向均存在. 此外假定存在这样一个方向，使得沿这个方向的导数是极小的. 这样的方向即称作泛函 Φ 在 x 点的最速下降方向.

---

*⁾ 两个非零元素 z_1 和 z_2 当且仅当它们成正比时（亦即存在某个 $\lambda>0$ 使得 $z_1=\lambda z_2$）或 z_1 和 z_2 落在从零点出发的同一射线上时才认为具有同一方向. 通常是用落在此射线上的某个向量来给出方向.

我们现在来描述最速下降法．假定泛函Φ在每一点处沿所有的方向均可微且对每一个$x\in\mathbf{X}$都存在最速下降方向．任取元素$x_0\in\mathbf{X}$（此元素叫极小元素的零阶逼近）．假定我们已经求出了k阶逼近x_k．下一次逼近x_{k+1}很自然地由形如

$$x_{k+1}=x_k+\varepsilon_{k+1}z_{k+1}$$

的方式来确定。

上式中的z_{k+1}是在点x_k处的最速下降方向．数值参数ε_{k+1}称作下降值，它在不同的情况下可以有种种不同的取法．这里只指出其中的三个求法：

1）设$\dfrac{\partial\Phi}{\partial z_{k+1}}(x)<0$而$\dfrac{\partial\Phi}{\partial z}(x_k+\alpha z_{k+1})$是$\alpha$的连续函数．这时函数$\varphi(\alpha;x,z)$至少在区间$[0,\alpha_k']$中下降，此处的$\alpha_k'$是方程$\varphi'\cdot(\alpha;x_k,z_{k+1})=0$的最小的正根．此时就取此根为下降值$\varepsilon_{k+1}$.

2）假定函数$\varphi(\alpha;x_k,z_{k+1})$在正半轴上某点处达到极小，就取这一点为下降值．

当Φ是严格凸泛函时，上面两种方法一致（详细的情形可见4.4的讨论）．

3）给定一正数序列$\{\varepsilon_k\}$使得$\varepsilon_k\to0$且$\sum\limits_{k=0}^{\infty}\varepsilon_k=\infty$（假定方向$z_k$已规格化，即$\|z_k\|=1$）．那么，我们就取第$k$步的下降值等于$\varepsilon_k$.

选定了求下降值确定的方法之后，相应地我们也就得到了序列$\{x_k\}$，在很多情况下它就是极小化序列．以上所说正是最速下降法的主要算法程式．

可用类似的算法来求上有界泛函的极大值．当然，此时最速下降方向应代之以最速上升的方向．

1.2. 求泛函的极小问题与解线性泛函方程问题之间可用下述方法来建立联系．

设U是 Hilbert 空间$\mathbf{H}$中的自共轭算子．象通常一样，我们总是把m和M分别表示为算子U的下界与上界（见 IX.3.4）．假定 $m>0$ 并考察方程

$$Ux=y. \tag{1}$$

因为 $\lambda=0$ 显然不属于算子U的谱，由定理 IX.5.3可知存在逆算子 U^{-1}，因而方程(1)对任何 $y\in\mathbf{H}$ 都存在唯一的解 $x^*=U^{-1}y$.

其次，我们作出如下的泛函

$$F(x)=(Ux,x)-[(x,y)+(y,x)] \quad (x\in\mathbf{H}). \tag{2}$$

定理 1. 方程(1)的解 x^* 使泛函(2)达到极小．反之，如果元素 $\tilde{x}$ 使泛函(2)达到极小，则 $\tilde{x}$ 是方程(1)的解，即 $\tilde{x}=x^*$.

证. 因 $y=Ux^*$，故泛函F可改写为

$$\begin{aligned} F(x)&=(Ux,x)-(x,Ux^*)-(Ux^*,x)\\ &=(U(x-x^*),x-x^*)-(Ux^*,x^*). \end{aligned} \tag{3}$$

由此

$$F(x)\geqslant m(x-x^*,x-x^*)-(Ux^*,x^*)\geqslant-(Ux^*,x^*)=F(x^*),$$

这就是说 x^* 使得泛函F达到极小.

由这个不等式还可以推出定理的第二部分．因为

$$\begin{aligned} 0=F(\tilde{x})-F(x^*)&=(U(\tilde{x}-x^*),\tilde{x}-x^*)\\ &\geqslant m(\tilde{x}-x^*,\tilde{x}-x^*), \end{aligned}$$

故有 $\tilde{x}=x^*$(注意 $m>0$).

注. 如果解 x^* 的存在是已知的，则由定理的证明可知，条件 $m>0$ 还可换成更弱的条件，即只要假定对任何非零元x有$(Ux,x)>0$即可.

1.3. 对泛函(2)应用最速下降法，就得到一个收敛于方程(1)的解 x^* 的序列．这就使得我们可将最速下降法看作是求方程(1)的近似解的一种方法.

我们现在来说明最速下降法应用于泛函F的格式.

如果 $x, z \in \mathbf{H}$, 则

$$F(x+z)=(U(x+z), x+z)-[(x+z, y)+(y, x+z)]$$
$$=(Ux, x)-[(x, y)+(y, x)]+[(Ux-y, z)$$
$$+(z, Ux-y)]+(Uz, z)$$
$$=F(x)+[(Ux-y, z)+(z, Ux-y)]-(Uz, z). \quad (4)$$

由此

$$\frac{\partial F}{\partial z}(x)=\frac{1}{\|z\|}[(Ux-y, z)+(z, Ux-y)]=\frac{2\mathrm{Re}(Ux-y, z)}{\|z\|}.$$

这样，在点 x_0 处的最速下降方向由元素 $(-z_1)$ 给出，其中

$$z_1=Ux_0-y.$$

为了求 ε_1，构造方程

$$\varphi'(\alpha; x_0, -z_1)=0,$$

由(4)，

$$\varphi(\alpha; x_0, -z_1)=F(x_0-\alpha z_1)=F(x_0)-2\alpha(z_1, z_1)+\alpha^2(Uz_1, z_1).$$

因此

$$\varepsilon_1=\frac{(z_1, z_1)}{(Uz_1, z_1)}.$$

最后得到

$$x_1=x_0-\varepsilon_1 z_1.$$

同样可得

$$x_2=x_1-\varepsilon_2 z_2 \qquad \left(z_2=Ux_1-y, \quad \varepsilon_2=\frac{(z_2, z_2)}{(Uz_2, z_2)}\right).$$

一般地，有

$$x_n=x_{n-1}-\varepsilon_n z_n \qquad (n=1, 2, \cdots),$$

其中

$$z_n=Ux_{n-1}-y, \quad \varepsilon_n=\frac{(z_n, z_n)}{(Uz_n, z_n)}, \quad (n=1, 2, \cdots).$$

上面所构造的 $\{x_n\}$ 是极小化序列，且有

定理 2. 序列$\{x_n\}$收敛于 x^*，收敛速度有如下估计

$$\|x_n - x^*\| \leqslant C\left(\frac{M-m}{M+m}\right)^n \qquad \left(C = \frac{\|z_1\|}{m};\ n = 0, 1, 2, \cdots\right). \tag{5}$$

证. 将方程(1)改写成

$$x = x - kUx + ky, \tag{6}$$

其中 $k>0$ 这样来选择，使得算子

$$T = I - kU$$

的范数尽可能小. 因算子T 的 上、下 界 分别 是$(1-km)$和$(1-kM)$，如果

$$1 - km = -(1-kM),$$

则范数的极小值将在

$$k = \frac{2}{M+m}$$

时达到. 这时

$$\|T\| = 1 - km = kM - 1 = \frac{M-m}{M+m}. \tag{7}$$

记

$$\tilde{x}_1 = Tx_0 + ky = x_0 - k(Ux_0 - y) = x_0 - kz_1. \tag{8}$$

再引进算子 $V = U^{1/2}$(见定理 V. 6. 2). 由(3)，可将泛函 F 写成

$$\begin{aligned} F(x) &= (U(x-x^*), x-x^*) - (Ux^*, x^*) \\ &= (V(x-x^*), V(x-x^*)) - (Vx^*, Vx^*) \\ &= \|V(x-x^*)\|^2 - \|Vx^*\|^2. \end{aligned} \tag{9}$$

但由显然成立的不等式 $F(\tilde{x}_1) \geqslant F(x_1)$ 可得

$$F(x_1) - F(x^*) \leqslant F(\tilde{x}_1) - F(x^*),$$

根据(9)，上不等式又可写成

$$\|V(x_1 - x^*)\| \leqslant \|V(\tilde{x}_1 - x^*)\|. \tag{10}$$

方程(6)等价于方程(1). 所以

$$x^* = Tx^* + ky.$$

从等式(8)减去上式，我们得到

$$\tilde{x}_1 - x^* = T(x_0 - x^*),$$

因而

$$V(\tilde{x}_1 - x^*) = TV(x_0 - x^*),$$

于是

$$\|V(\tilde{x}_1 - x^*)\| \leqslant \|T\| \|V(x_0 - x^*)\| = \frac{M-m}{M+m} \|V(x_0 - x^*)\|.$$

将上式与(10)式比较，得

$$\|V(x_1 - x^*)\| \leqslant \frac{M-m}{M+m} \|V(x_0 - x^*)\|. \tag{11}$$

对每个 $n=1, 2, \cdots$ 作类似的论证，可得

$$\|V(x_n - x^*)\| \leqslant \frac{M-m}{M+m} \|V(x_{n-1} - x^*)\| \qquad (n=1, 2, \cdots).$$

最后得到

$$\|V(x_n - x^*)\| \leqslant \left(\frac{M-m}{M+m}\right)^n \|V(x_0 - x^*)\| \quad (n=1, 2, \cdots). \tag{12}$$

因为函数 $t^{-1/2}$ 在$[m, M]$上连续，它当然更在算子U的谱上连续，所以存在逆算子 V^{-1}. 并且有(见定理 IX. 5.2)

$$\|V^{-1}\| = \max_{t \in S_U} \frac{1}{\sqrt{t}} = \frac{1}{\sqrt{m}}.$$

我们还指出

$$\|V\| = \max_{t \in S_U} \sqrt{t} = \sqrt{M}.$$

根据以上的讨论，利用(12)后得

$$\|x_n - x^*\| = \|V^{-1}V(x_n - x^*)\| \leqslant \|V^{-1}\| \left(\frac{M-m}{M+m}\right)^n \|V(x_0 - x^*)\|$$

$$\leqslant \|V^{-1}\| \|V\| \left(\frac{M-m}{M+m}\right)^n \|x_0 - x^*\| = \sqrt{\frac{M}{m}} \left(\frac{M-m}{M+m}\right)^n \|x_0 - x^*\|.$$

这个估计式不足的一点是它的右端含有未知元素 x^*. 为了避免出现它，注意有

$$\|V(x_0-x^*)\|=\|V^{-1}U(x_0-x^*)\|\leqslant\|V^{-1}\|\|Ux_0-y\|$$

$$=\frac{\|z_1\|}{\sqrt{m}}.$$

由此可得

$$\|x_n-x^*\|\leqslant\frac{\|z_1\|}{m}\left(\frac{M-m}{M+m}\right)^n,\qquad (n=1,2,\cdots),$$

这和(5)式一致.

定理证毕.

1.4. 现在用 x_0 和 z_1 来表示 p 阶近似 x_p. 当 $p=2$ 时，我们有

$$x_2=x_1-\varepsilon_2z_2=x_1-\varepsilon_2(Ux_1-y)$$

$$=x_0-\varepsilon_1z_1-\varepsilon_2(Ux_0-y-\varepsilon_1Uz_1)$$

$$=x_0-(\varepsilon_1+\varepsilon_2)z_1+\varepsilon_1\varepsilon_2Uz_1$$

$$=x_0+\lambda_1^{(2)}z_1+\lambda_2^{(2)}Uz_1.$$

按归纳法不难证明

$$x_p=x_0+\lambda_1^{(p)}z_1+\cdots+\lambda_p^{(p)}U^{p-1}z_1. \tag{13}$$

考察元素

$$x_{(1)}=x_0+\lambda_1z_1+\cdots+\lambda_pU^{p-1}z_1,$$

并按使泛函 $F(x_{(1)})$ 达到极小的要求来确定诸复系数 $\lambda_1,\lambda_2,\cdots,\lambda_p$. 因为

$$F(x_{(1)})=\left(Ux_0+\sum_{j=1}^{p}\lambda_jU^jz_1,\ \ x_0+\sum_{j=1}^{p}\lambda_jU^{j-1}z_1\right)$$

$$-[(x_{(1)},y)+(y,x_{(1)})]$$

$$=F(x_0)+\sum_{j=1}^{p}(\lambda_j+\bar{\lambda}_j)(U^{j-1}z_1,z_1)+$$

$$+\sum_{j=1}^{p}\sum_{k=1}^{p}\lambda_j\bar{\lambda}_k(U^j z_1, U^{k-1}z_1),$$

如果记 $\lambda_k=\sigma_k+\tau_k i\,(k=1, 2, \cdots, p)$，则我们就得到确定 λ_k 的两组等式

$$\frac{\partial F(x_{(1)})}{\partial \sigma_j}=0, \quad \frac{\partial F(x_{(1)})}{\partial \tau_j}=0 \qquad (j=1, 2, \cdots, p).$$

将这两组等式的左端具体写出来，就是

$$(U^{j-1}z_1, z_1)+\sum_{k=1}^{p}\sigma_k(U^j z_1, U^{k-1}z_1)=0,$$

$$\sum_{k=1}^{p}\tau_k(U^j z_1, \ U^{k-1}z_1)=0.$$

后一等式乘上 i 并和前一个等式相加，并注意到 $(U^j z_1, U^{k-1}z_1)=(U^{j+k-1}z_1, z_1)$，则可将这两组等式合并写成一组：

$$(U^{j-1}z_1, z_1)+\sum_{k=1}^{p}\lambda_k(U^{j+k-1}z_1, z_1)=0$$

$$(j=1, 2, \cdots, p). \tag{14}$$

由于所求的极小显然存在，故(14)必有解*).

由元素 $x_{(1)}$ 出发，用类似的办法构造出 $x_{(2)}$，如此等等，最后即得到一个序列 $x_{(0)}=x_0, x_{(1)}, \cdots$，此序列对应于最速下降法的 p 步变形.

从这个序列的构造方法显然可见它收敛于 x^* 的速度至少比序列 $\{x_n\}$ 快 p 倍. 事实上，利用定理 2 的证明不难得到

$$\|x_{(n)}-x^*\|\leqslant\frac{\|z_1\|}{m}\left[\left(\frac{M-m}{M+m}\right)^p\right]^n. \tag{15}$$

然而应该期望实际的收敛速度还要快一些.

---

*) 尽管这种解可能不唯一. 但应指出，值 $F(x_{(1)})$ 与解的选择没有关系.

定理 3(Бирман). 成立如下的估计式

$$\|x_{(n)}-x^*\|\leqslant\frac{\|z_1\|}{m}\left[\frac{1}{\theta_p\left(\frac{M+m}{M-m}\right)}\right]^n,$$

$$(n=1,2,\cdots),\tag{16}$$

其中 $\theta_p(t)$是 p 次 Чебышёв 多项式

$$\theta_p(t)=\cos p\arccos t\qquad (t\in[-1,1]).$$

证. 考察方程

$$x=Tx+\tilde{y}\qquad (T=I-U\varphi(U),\tilde{y}=\varphi(U)y),\tag{17}$$

其中 $\varphi(t)$是 $p-1$ 次多项式.

命

$$\tilde{x}_1=Tx_0+\tilde{y}=x_0-\varphi(U)(Ux_0-y)=x_0-\varphi(U)z_1$$

$$=x_0+\sum_{k=1}^{p}c_kU^{k-1}z_1.$$

显然有 $F(\tilde{x}_1)\geqslant F(x_{(1)})$, 这时重复定理 2 的讨论可得

$$\|x_{(n)}-x^*\|\leqslant\frac{\|z_1\|}{m}\|T\|^n\qquad (n=1,2,\cdots).\tag{18}$$

因为

$$\|T\|\leqslant\max_{t\in[m,M]}|1-t\varphi(t)|,$$

所以自然用使得 $\max\limits_{t\in[m,M]}|1-t\varphi(t)|$ 为极小的条件来确定多项式 $\varphi(t)$, 如果记 $\psi(t)=1-t\varphi(t)$, 则 $\psi(t)$是在区间$[m, M]$上和零偏差最小的 p 次多项式, 且 $\psi(0)=1$. 大家知道*),

$$\psi(t)=\frac{\theta_p\left(\frac{2t-M-m}{M-m}\right)}{\theta_p\left(\frac{M+m}{M-m}\right)}.$$

---

*) 见 Гончаров[1], 第 230 页.

因为

$$\max_{t\in[m,M]}\left|\theta_p\left(\frac{2t-M-m}{M-m}\right)\right|=\max_{t\in[-1,1]}|\theta_p(t)|=1,$$

故得对$\|T\|$的估计

$$\|T\|\leqslant\max_{t\in[m,M]}|\psi(t)|=\frac{1}{\theta_p\left(\frac{M+m}{M-m}\right)},$$

再由(18)即证得了定理.

注 1. 设 $\rho(t)$是任意的p次多项式，这时成立如下的不等式

$$|\rho(t_0)|\leqslant|\theta_p(t_0)|\max_{t\in[-1,1]}|\rho(t)|\qquad(|t_0|>1).$$

如果取

$$\rho(t)=t^p,\quad t_0=\frac{M+m}{M-m}$$

则有

$$\frac{1}{\theta_p\left(\frac{M+m}{M-m}\right)}\leqslant\left(\frac{M-m}{M+m}\right)^p.$$

因此，(15)式乃是估计式(16)的推论.

注 2. 因为

$$\frac{\theta_p(t)}{t^p}\xrightarrow[t\to\infty]{}\frac{1}{2^{p-1}},$$

如果$\frac{M+m}{M-m}$充分大，则(16)与(15)的差别是一个趋于$\left(\frac{1}{2^{p-1}}\right)^n$的因子.

在 Самокиш[1]中，在对算子U的谱作若干补充假定之下得到了比(16)更精确的估计.

注 3. 当 $m=0$ 时可以证明，如果方程(1)的解存在(不一定唯一)，则最速下降法收敛于接近于x_0 的解 x^*(Фридман[1]). 并且有估计式 $F(x_n)-F(x^*)=O(1/n)$，但 x_n 收敛于 x^* 的速度可能相

当慢.

§2. 求紧算子的特征值

2.1. 现在我们来考察紧自共轭算子 U. 不失一般性, 可设 $M > |m| \geqslant 0$(参见 IX. 4.3). 于是就有

$$M = \sup_{\|x\|=1}(Ux, x) = \sup_{x\neq 0}\frac{(Ux, x)}{(x, x)} = \frac{(Ux^*, x^*)}{(x^*, x^*)},$$

其中 x^* 是对应于最大特征值 $\lambda_1 = M$ 的特征元素. 因此, λ_1 是泛函

$$L(x) = \frac{(Ux, x)}{(x, x)}$$

的最大值, 由 IX. 4.3 所述, 每个使该泛函达到极大的元素都是特征元素.

为了求泛函 L 的极大值, 我们应用最速下降法(确切地说应该是采用"最速上升法"). 取任意规格化元素 $x_0 \in \mathbf{H}$, 因为

$$L(x_0 + \alpha z) = \frac{(Ux_0, x_0) + \alpha(Ux_0, z) + \alpha(z, Ux_0) + \alpha^2(Uz, z)}{1 + \alpha(x_0, z) + \alpha(z, x_0) + \alpha^2(z, z)}, \tag{1}$$

所以

$$\begin{aligned}&\frac{\partial L}{\partial z}(x_0)\\ &= \frac{[(Ux_0, z) + (z, Ux_0)] - [(x_0, z) + (z, x_0)](Ux_0, x_0)}{\|z\|}\\ &= \frac{(Ux_0 - \mu_0 x_0, z) + (z, Ux_0 - \mu_0 x_0)}{\|z\|},\end{aligned}$$

其中

$$\mu_0 = L(x_0) = (Ux_0, x_0).$$

由表达式$\frac{\partial L}{\partial z}(x_0)$可见"梯度"方向由元素 $z_1 = Ux_0 - \mu_0 x_0$ 给出(假

定 $z_1\neq 0$).

于是，ε_1 由方程

$$\varphi'(\alpha; x_0, z_1)=0 \tag{2}$$

确定，其中

$$\varphi(\alpha; x_0, z_1)=L(x_0+\alpha z_1).$$

由于

$$(x_0, z_1)=(x_0, Ux_0-\mu x_0)=(x_0, Ux_0)-\mu_0(x_0, x_0)=0,$$

$$(Ux_0, z_1)=(\mu_0 x_0+z_1, z_1)=\mu_0(x_0, z_1)+(z_1, z_1)$$
$$=(z_1, z_1),$$

经过一些初等变形，方程(2)可写成

$$\frac{2\{(z_1, z_1)+[(Uz_1, z_1)-\mu_0(z_1, z_1)]\alpha-(z_1, z_1)^2\alpha^2\}}{[1+\alpha^2(z_1, z_1)]^2}=0.$$

于是，ε_1 就是方程

$$(z_1, z_1)^2\alpha^2-[(Uz_1, z_1)-\mu_0(z_1, z_1)]\alpha-(z_1, z_1)=0$$

的正根，亦即

$$\varepsilon_1=\frac{[(Uz_1, z_1)-\mu_0(z_1, z_1)]+\sqrt{[(Uz_1, z_1)-\mu_0(z_1, z_1)]^2+4(z_1, z_1)^3}}{2(z_1, z_1)^2}.$$

命

$$x^{(1)}=x_0+\varepsilon_1 z_1,$$

我们取元素

$$x_1=\frac{x^{(1)}}{\|x^{(1)}\|}=\frac{x_0+\varepsilon_1 z_1}{\|x_0+\varepsilon_1 z_1\|}$$

作为下一次近似.

用 x_1 去取代 $x^{(1)}$ 的地位并继续上述过程，我们就得到序列 x_0, $x_1, x_2, \cdots, x_n, \cdots$

$$x_n=\frac{x^{(n)}}{\|x^{(n)}\|}, \quad x^{(n)}=x_{n-1}+\varepsilon_n z_n \qquad (n=1, 2, \cdots),$$

其中

$$z_n = Ux_{n-1} - \mu_{n-1}x_{n-1}, \quad \mu_{n-1} = L(x_{n-1}) = (Ux_{n-1}, x_{n-1}),$$

$$\varepsilon_n = \frac{[(Uz_n, z_n) - \mu_{n-1}(z_n, z_n)] + \sqrt{[(Uz_n, z_n) - \mu_{n-1}(z_n, z_n)]^2 + 4(z_n, z_n)^3}}{2(z_n, z_n)^2}.$$

2.2. 对于序列$\{\mu_n\}$收敛于λ_1以及序列$\{x_n\}$收敛于对应于λ_1的特征元素的问题，下面的定理作出了回答.

定理 1. 如果x_0与对应于最大特征值$\lambda_1 = M$的特征子空间不正交，则$\mu_n \to \lambda_1$，$x_n \to x^*$，其中x^*是属于特征值λ_1的特征元素.

证. 用$\lambda_k (k=1, 2, \cdots)$表示算子$U$的各个特征值，用$\mathbf{H}_{\lambda_k} = \mathbf{H}_k$表示对应于$\lambda_k$的特征子空间，用$P_k$表示作用于这些子空间上的投影算子. 其次，命

$$x_k^* = \frac{P_k x_0}{\|P_k x_0\|}\ ^{*)}.$$

根据 IX.4.5 有

$$x_0 = \sum_k P_k x_0 = \sum_k \|P_k x_0\| x_k^* = \sum_k c_{0k} x_k^*$$

$$(c_{0k} = \|P_k x_0\|, \quad k = 1, 2, \cdots). \tag{3}$$

易见$c_{0k} \geqslant 0 (k=1, 2, \cdots)$，由定理的条件$c_{01} > 0$. 由于(3)式右端各加项是两两正交的，又$\|x_k^*\| = 1$，所以

$$\sum_k c_{0k}^2 = \|x_0\|^2 = 1,$$

于是，最后得

$$\mu_0 = (Ux_0, x_0) = \sum_k \lambda_k c_{0k}^2 \qquad (\mu_0 \leqslant \lambda_1).$$

元素x_1也可用$x_1^*, x_2^*, \cdots$来表示. 事实上，

$$x_1 = \frac{x^{(1)}}{\|x^{(1)}\|} = \frac{x_0 + \varepsilon_1 z_1}{\|x^{(1)}\|} = \frac{x_0 + \varepsilon_1(Ux_0 - \mu_0 x_0)}{\|x^{(1)}\|}.$$

---

*) 如果$P_k x_0 = 0$，则$x_k^* = 0$.

又因

$$Ux_0=\sum_k \lambda_k c_{0k} x_k^*,$$

所以

$$x_1=\sum_k c_{1k} x_k^*$$

$$\left(c_{1k}=\frac{[1+\varepsilon_1(\lambda_k-\mu_0)]c_{0k}}{\|x^{(1)}\|};\ k=1,2,\cdots\right).$$

和上面一样

$$\sum_k c_{1k}^2=1 \qquad \mu_1=\sum_k \lambda_k c_{1k}^2 \qquad c_{11}>0.$$

一般地

$$x_n=\sum_k c_{nk} x_k^* \tag{4}$$

$$\left(c_{nk}=\frac{[1+\varepsilon_n(\lambda_k-\mu_{n-1})]c_{n-1k}}{\|x^{(n)}\|};\ \ k,n=1,2,\cdots\right),$$

并且

$$\sum_k c_{nk}^2=1,\ \ \mu_n=\sum_k \lambda_k c_{nk}^2,\ \ c_{n1}>0\ \ (n=0,1,2,\cdots). \tag{5}$$

现在来找出用 μ_{n-1} 表示 μ_n 的关系式. 对应于公式(1)有

$$\mu_n=L(x_n)=L(x^{(n)})=L(x_{n-1}+\varepsilon_n z_n)$$
$$=\frac{(Ux_{n-1},x_{n-1})+2\varepsilon_n(Ux_{n-1},z_n)+\varepsilon_n^2(Uz_n,z_n)}{1+2\varepsilon_n(x_{n-1},z_n)+\varepsilon_n^2(z_n,z_n)}.$$

但是

$$(x_{n-1},z_n)=(x_{n-1},Ux_{n-1}-\mu_{n-1}x_{n-1})=0,$$
$$(Ux_{n-1},z_n)=(\mu_{n-1}x_{n-1}+z_n,z_n)=(z_n,z_n).$$

因此

$$\mu_n=\frac{\mu_{n-1}+2\varepsilon_n(z_n,z_n)+\varepsilon_n^2(Uz_n,z_n)}{1+\varepsilon_n^2(z_n,z_n)}=$$

$$= \mu_{n-1} + \frac{2\varepsilon_n(z_n, z_n) + \varepsilon_n^2[(Uz_n, z_n) - \mu_{n-1}(z_n, z_n)]}{1 + \varepsilon_n^2(z_n, z_n)}.$$

同时，ε_n 满足方程

$$(z_n, z_n)^2\varepsilon_n^2 - [(Uz_n, z_n) - \mu_{n-1}(z_n, z_n)]\varepsilon_n - (z_n, z_n) = 0, \quad (6)$$

于是就有

$$\varepsilon_n^2[(Uz_n, z_n) - \mu_{n-1}(z_n, z_n)] = \varepsilon_n^3(z_n, z_n)^2 - \varepsilon_n(z_n, z_n).$$

由此，我们得到

$$\mu_n = \mu_{n-1} + \frac{2\varepsilon_n(z_n, z_n) + \varepsilon_n^3(z_n, z_n)^2 - \varepsilon_n(z_n, z_n)}{1 + \varepsilon_n^2(z_n, z_n)},$$

化简，最后得到

$$\mu_n = \mu_{n-1} + \varepsilon_n(z_n, z_n). \quad (7)$$

因 $\varepsilon_n > 0$，故 $\mu_n > \mu_{n-1}$，所以由序列$\{\mu_n\}$的上有界性($\mu_n \leqslant \lambda_1$)可知极限存在，我们把它记为 μ，即

$$\mu = \lim_{n\to\infty} \mu_n.$$

由(7)推知

$$\lim_{n\to\infty} \varepsilon_n(z_n, z_n) = 0, \quad (8)$$

另外，由

$$|\varepsilon_n(Uz_n, z_n)| \leqslant \|U\| \varepsilon_n(z_n, z_n)$$

还得到

$$\lim_{n\to\infty} \varepsilon_n(Uz_n, z_n) = 0. \quad (9)$$

由(8)，(9)，在方程(6)中让 $n\to\infty$ 而取极限，就得到

$$\lim_{n\to\infty}(z_n, z_n) = 0.$$

因而

$$\lim_{n\to\infty}[Ux_n - \mu_n x_n] = \lim_{n\to\infty} z_{n+1} = 0. \quad (10)$$

由于 U 是紧算子而$\{x_n\}$是有界序列，所以存在子序列$\{x_{n_j}\}$使

得$\{Ux_{n_j}\}$收敛．从关系式(10)推知序列$\{x_{n_j}\}$收敛．设$x_{n_j}\to\tilde{x}$．在等式

$$Ux_{n_j}-\mu_{n_j}x_{n_j}=z_{n_{j+1}}$$

中取极限后得到

$$U\tilde{x}-\mu\tilde{x}=\mathbf{0}.$$

因此，μ是算子U的特征值而$\tilde{x}$是相应的特征元素．现在来证明$\mu=\lambda$而且$\tilde{x}=x_1^*$．设$\mu=\lambda_s(s>1)$．由(4)可知

$$\tilde{x}=\sum_k c_k x_k^* \qquad (c_k=\lim_{j\to\infty}c_{n_jk}\,;\ k=1,2,\cdots).$$

但$k\neq s$时$(\tilde{x},x_k^*)=0$，因而$\tilde{x}=c_s x_s^*$且$|c_s|=1$．由于$c_{ns}\geqslant 0$，所以最后得知$c_s=1$而$\tilde{x}=x_s^*$．

考虑比式$\dfrac{c_{ns}}{c_{n1}}$，由(4)式，得到

$$\frac{c_{ns}}{c_{n1}}=\frac{1+\varepsilon_n(\lambda_s-\mu_{n-1})}{1+\varepsilon_n(\lambda_1-\mu_{n-1})}\,\frac{c_{n-1,s}}{c_{n-1,1}}<\frac{c_{n-1,s}}{c_{n-1,1}},$$

亦即这个比式随n的增大而减小．但$\lim\limits_{j\to\infty}c_{n_js}=c_s=1$，故$\lim\limits_{j\to\infty}c_{n_j1}=c_1=0$，因此

$$\lim_{j\to\infty}\frac{c_{n_js}}{c_{n_j1}}=\infty,$$

但这和前面所建立的结果相矛盾，因而$\mu=\lambda_1$，$\tilde{x}=x_1^*$．

现在来证明整个序列$\{x_n\}$收敛于x_1^*．设d是从点λ_1到算子U的谱中去掉λ_1后所余部分的距离．我们就有

$$\begin{aligned}\|x_n-x_1^*\|^2&=\|x_n\|^2+\|x_1^*\|^2-(x_n,x_1^*)-(x_1^*,x_n)\\&=2(1-c_{n1})\leqslant 2(1-c_{n1}^2)=2\sum_{k\geqslant 2}c_{nk}^2\\&\leqslant\frac{2}{d}\sum_k(\lambda_1-\lambda_k)c_{nk}^2.\end{aligned}$$

注意到(5)式，则

$$\sum_k (\lambda_1-\lambda_k)c_{nk}^2=\lambda_1\sum_k c_{nk}^2-\sum_k \lambda_k c_{nk}^2=\lambda_1-\mu_n,$$

于是

$$\|x_n-x_1^*\|^2\leqslant\frac{2}{d}(\lambda_1-\mu_n)\qquad(n=0,1,2,\cdots).\tag{11}$$

因为 $\mu_n\to\lambda_1$，所以由(11)式可得 $x_n\to x_1^*$.

定理全部证毕.

注. 关于 $\mu_n\to\lambda_1$ 的收敛速度，有如下估计

$$\lambda_1-\mu_n\leqslant q^2(1+\alpha_n)(\lambda_1-\mu_{n-1})\qquad(n=1,2,\cdots),$$

其中

$$q=\frac{M-m-d}{M-m+d},$$

而序列 $\{\alpha_n\}$ 单调下降而趋于零.

因此，如果 n 充分大，使得 $q(1+\alpha_n)<1$，则收敛速度相当于几何级数的收敛速度.

由估计式(11)还可以刻画 $\{x_n\}$ 收敛于 x_1^* 的速度.

关于用最速下降法求特征值的讨论可参见 Самокиш[2].

§3. 对椭圆型微分方程的应用

3.1. 我们来考察应用最速下降法解椭圆型微分方程的问题. 为简单计，只考虑具有两个变元的二阶自共轭方程的边值问题:

$$Lx\equiv-\frac{\partial}{\partial s}\left(a\frac{\partial x}{\partial s}\right)-\frac{\partial}{\partial t}\left(b\frac{\partial x}{\partial t}\right)+cx=\varphi,\quad x|_S=0,\tag{1}$$

其中 $x\in D$，而 D 是由中心在坐标原点的圆周 S 所界的区域，系数 a 和 b 假定是连续可微函数，c 和 φ 是连续函数，此外，还设在 D 内有 $a,b>0$, $c\geqslant0$.

方程(1)两端作用以算子 $-\Delta^{-1}$，得到一个新的方程

$$Ux\equiv-\Delta^{-1}Lx=\psi\qquad(\psi=-\Delta^{-1}\varphi),\tag{2}$$

我们将在子空间 $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ 中(即由 $\mathbf{W}_2^{(1)}(D)$ 中在 S 上取值为零的函数所构成的空间)来解这个方程．正象在 XIV. 6. 3 中所指出的那样，Δ^{-1} 是从 $\mathbf{L}^2$ 到 $\overset{\circ}{\mathbf{W}}{}_2^{(2)}$ 的连续线性算子，于是 $\psi\in\overset{\circ}{\mathbf{W}}{}_2^{(2)}$.

我们来验证，算子 $U=-\Delta^{-1}L$ 满足 §1 中的条件．设 x 是 $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ 中二次连续可微函数．因 Lx 也是连续函数，所以 $y=Ux=-\Delta^{-1}Lx\in\overset{\circ}{\mathbf{W}}{}_2^{(2)}\subset\overset{\circ}{\mathbf{W}}{}_2^{(1)}$．应用 Green 公式两次即得

$$
\begin{aligned}
(Ux,x)&=(-\Delta^{-1}Lx,x)=(y,x)=\iint_D\left(\frac{\partial y}{\partial s}\frac{\partial x}{\partial s}+\frac{\partial y}{\partial t}\frac{\partial x}{\partial t}\right)dsdt\\
&=-\iint_D x\left[\frac{\partial^2 y}{\partial s^2}+\frac{\partial^2 y}{\partial t^2}\right]dsdt=-\iint_D x\Delta y\,dsdt\\
&=\iint_D x\left[-\frac{\partial}{\partial s}\left(a\frac{\partial x}{\partial s}\right)-\frac{\partial}{\partial t}\left(b\frac{\partial x}{\partial t}\right)+cx\right]dsdt\\
&=\iint_D\left[a\left(\frac{\partial x}{\partial s}\right)^2+b\left(\frac{\partial x}{\partial t}\right)^2+cx^2\right]dsdt. \qquad (3)
\end{aligned}
$$

由这个关系式给出

$$
(Ux,x)\geqslant\alpha\iint_D\left[\left(\frac{\partial x}{\partial s}\right)^2+\left(\frac{\partial x}{\partial t}\right)^2\right]dsdt=\alpha(x,x), \qquad (4)
$$

其中

$$
\alpha=\min[\min a(s,t),\ \min b(s,t)] \qquad ((s,t)\in D).
$$

根据嵌入定理(即定理 XI. 4. 3)

$$
\iint_D|x(s,t)|^2dsdt\leqslant A\|x\|^2_{\overset{\circ}{\mathbf{W}}{}_2^{(1)}},
$$

将这个结果应用于(3)式，则有

$$
(Ux,x)\leqslant\beta\iint_D\left[\left(\frac{\partial x}{\partial s}\right)^2+\left(\frac{\partial x}{\partial t}\right)^2\right]dsdt=\beta(x,x), \qquad (5)
$$

其中

$$\beta=2\max[\max a(s,t),\max b(s,t),A\max c(s,t)]$$
$$((s,t)\in D).$$

其次，象导出(3)那样，我们可推知等式

$$(Ux,y)=(x,Uy) \tag{6}$$

对任何二次连续可微函数 x, $y\in\mathring{W}_2^{(1)}$ 均成立. 对照(4)、(5)和(6)可知作为从 $\mathring{W}_2^{(1)}$ 到二次连续可微函数集合上的算子 U 是有界的，又因二次连续可微函数的集合在 $\mathring{W}_2^{(1)}$ 中稠密，所以 U 也是 $\mathring{W}_2^{(1)}$ 上的有界算子. 关系式(4)、(5)和(6)这时可延拓到整个空间 $\mathring{W}_2^{(1)}$ 上. 因此 U 是自共轭算子并且它的界 m 和 M 有如下估计

$$m\geqslant\alpha>0,\qquad M\leqslant\beta. \tag{7}$$

我们现在对方程(2)应用最速下降法. 如果 $x_0\in\mathring{W}_2^{(1)}$ 是初次逼近，则逼近序列 x_n 由可按 XV.1.3 中的公式

$$x_n=x_{n-1}-\varepsilon_n z_n\qquad(n=1,2,\cdots)$$

决定，其中

$$z_n=Ux_{n-1}-\psi=\Delta^{-1}(Lx_{n-1}-\varphi),$$

亦即，z_n 由方程

$$\Delta z=Lx_{n-1}-\varphi,\quad z|_S=0$$

求出. 至于 ε_n，由(3)可知

$$\varepsilon_n=\frac{(z_n,z_n)}{(Uz_n,z_n)}=\frac{\iint\limits_D\left[\left(\frac{\partial z_n}{\partial s}\right)^2+\left(\frac{\partial z_n}{\partial t}\right)^2\right]dsdt}{\iint\limits_D\left[a\left(\frac{\partial z_n}{\partial s}\right)^2+b\left(\frac{\partial z_n}{\partial t}\right)^2+cz_n^2\right]dsdt}.$$

关于这个过程的收敛性，由定理 1.2 及式(7)，得

$$\|x_n-x^*\|\leqslant C\left(\frac{M-m}{M+m}\right)^n\leqslant C\left(\frac{\beta-\alpha}{\beta+\alpha}\right)^n$$
$$(n=0,1,2,\cdots), \tag{8}$$

其中 x^* 是方程(2)的解，因而也是边值问题(1)的解.

3.2. 如果方程(1)中的系数 a,b,c 以及右端 φ 是多项式，则

如上所构造的逐次逼近过程特别有效. 这时,初值 x_0 也取为多项式. 为确定 z_1,我们得到方程

$$\Delta z = p, \tag{9}$$

其中 p 是多项式. 不难看出,这个方程在 $\overset{\circ}{W}{}_2^{(1)}$ 中的解仍然是多项式,其次数比多项式 p 的次数高两次. 事实上,将 z 写成

$$z(s, t) = (s^2 + t^2 - 1)\pi(s, t), \tag{10}$$

其中 π 是一待定的多项式,其次数与多项式 p 的次数相等.将(10)式代入方程(9)并让左、右两端的系数相等, 我们就得到确定多项式 π 的系数的一个线性代数方程组, 而且这个方程组中自变量的个数与方程的个数相等. 这时, 该方程组的系数矩阵显然不依赖于 p. 由于方程(9)的多项式解(如果存在的话)是唯一的, 所以这个方程组的系数行列式异于零,这就保证了对任何多项式 p, 方程(9)存在多项式解*).

这就说明 z_1 是多项式, 从而 $x_1 = x_0 - \varepsilon_1 z_1$ 也是多项式. 容易看到所有的各次逼近 x_n 都是多项式.

在对方程(1)作出熟知的正规性假定的前提下,上面所指出的情况不仅可建立方程(1)的广义解的存在性而且还可以建立它的古典意义下的解的存在性. 为此, 我们还需要一些函数构造论方面的知识.

3.3. 首先指出两个类似于 Бернштейн 和 Марков 不等式的不等式. 设 p 是 n 次多项式,并且

$$|p(s, t)| \leqslant M \qquad (s, t \in D),$$

这里 D 是界于光滑曲线的任意有界域. 这时,在区域 D 中有

$$\left|\frac{\partial p(s, t)}{\partial s}\right| \leqslant An^2 M, \quad \left|\frac{\partial p(s, t)}{\partial t}\right| \leqslant An^2 M, \tag{11}$$

*) 对于方程(9)的实际求解来说,比较简单的方法是: 先不管边值条件而求解,然后再求具有零边值问题的解,减去与前面求得的多项式边界条件相同的 Laplace 方程的解. 后一解若借助于将解展成 Fowrier 级数来求则非常简单.

而在D的任意内子域 D_1 中则有

$$\left|\frac{\partial p(s, t)}{\partial s}\right| \leqslant BnM, \quad \left|\frac{\partial p(s, t)}{\partial t}\right| \leqslant BnM, \qquad (12)$$

其中A和B是只依赖于区域D和D_1的常数.

不等式(11)和(12)可很显然地从相应的一维不等式得到(参见 Натансон-I, 第 169 页和 174 页).

我们还要证明如下引理.

引理 1. 如果u是次数不超过n次的多项式, 则

$$\|u\|_{C(D)} \leqslant A_1 n^2 \|u\|_{L^2(D)}, \quad \|u\|_{C(D_1)} \leqslant B_1 n \|u\|_{L^2(D)}, \qquad (13)$$

$$\|u\|_{C^{(1)}(D)} \leqslant A_2 n^2 \|u\|_{W_2^{(1)}(D)}, \quad \|u\|_{C^{(1)}(D_1)} \leqslant B_1 n \|u\|_{W_2^{(1)}(D)}, \qquad (14)$$

其中 D_1 表示D的内子域而 $C^{(1)}$表示连续可微函数空间.

证. 我们只限于D是中心在坐标原点的圆的情形来证明(当D是一般的有界域时, 只要作不大的补充论证即可). 命

$$\varphi(s, t) = \int_0^s \int_0^t p^2(s, t) ds dt,$$

我们就有

$$|\varphi(s, t)| \leqslant \|p\|_{L^2}^2, \quad ((s, t) \in D),$$

两次应用不等式(11)后得

$$\begin{aligned} |p(s, t)|^2 &= \left|\frac{\partial^2 \varphi(s, t)}{\partial s \partial t}\right| \\ &\leqslant A^2 (2n+2)^2 (2n+1)^2 \max|\varphi(s, t)| \\ &\leqslant A_1^2 n^4 \|p\|_{L^2}^2, \end{aligned}$$

由此即得(13)中的第一个不等式. 用(12)式替换(11)式, 可类似地建立(13)中的第二个不等式.

将刚建立的不等式(13)应用于导数, 则

$$\left|\frac{\partial p(s, t)}{\partial s}\right| \leqslant A_1 n^2 \left\|\frac{\partial p}{\partial s}\right\|_{L^2} \leqslant A_1 n^2 \|p\|_{W_2^{(1)}};$$

对导数$\frac{\partial p}{\partial t}$也成立类似的不等式. 最后,由(11)并根据嵌入定理, 可得

$$|p(s, t)| \leqslant An^2 \|p\|_{L^2} \leqslant CAn^2 \|p\|_{W_2^{(1)}}.$$

对比所有这些不等式, 我们就得到(14)中的第一个不等式

$$\|p\|_{C^{(1)}(D)} = \max_D |p(s, t)| + \max_D \left|\frac{\partial p(s, t)}{\partial s}\right| + \max_D \left|\frac{\partial p(s, t)}{\partial t}\right|$$
$$\leqslant A_2 n^2 \|p\|_{W_2^{(1)}}.$$

(14)中的第二个不等式可类似地得到.

3.4. 我们再回到方程(1)来讨论. 先设系数 a, b, c 和右端 φ 都是多项式. 如果初值 x_0 也是多项式, 则正如 3.2 中已经讲过的那样, 则所有以后的各次逼近 $x_n(n=1, 2, \cdots)$ 也都是多项式. 这时, 如果 a, b, c 和 φ 是次数不超过m 次的多项式, 而 x_0 是N次多项式, 则 Lx_0 将是次数不超过 $m+N$ 次的多项式, z_1将是次数不超过 $m+N+2$ 次的多项式. 特别, 如果 $x_0=0$, 则 z_1 因而 x_1 是次数不超过 $m+2$ 次的多项式. 重复上述的讨论, 即可得知 x_n 是次数不超过$(m+2)n$ 次的多项式.

根据估计式(8), 我们得到

$$\|x_k - x_{k-1}\|_{\mathring{W}_2^{(1)}} \leqslant \|x_k - x^*\| + \|x_{k-1} - x^*\|$$
$$\leqslant 2C\left(\frac{\beta-\alpha}{\beta+\alpha}\right)^{k-1}$$
$$(k=1, 2, \cdots).$$

因为差 $x_k - x_{k-1}$ 是次数不超过$(m+2)k$次的多项式, 所以由引理可得

$$\|x_k - x_{k-1}\|_{C^{(1)}(D)} \leqslant A_3(m+2)^2 k^n q \qquad (k=1, 2, \cdots),$$

其中

$$q = \frac{\beta-\alpha}{\beta+\alpha}.$$

由此, 多次应用不等式(11)后得

$$\|x_k - x_{k-1}\|_{C^{(p)}(D)} \leqslant A_4[(m+2)k]^{2p} q^k \qquad (k, p = 1, 2, \cdots). \tag{15}$$

因而,级数

$$x^*(s, t) = \sum_{k=1}^{\infty} [x_k(s, t) - x_{k-1}(s, t)] \tag{16}$$

可以逐项微分 p 次. 换言之, 方程(1)的解具有任意阶的导数, 而且

$$\|x^* - x_n\|_{C^{(p)}(D)} = \sum_{k=n+1}^{\infty} \|x_k - x_{k+1}\| \leqslant A_5[(m+2)n]^{2p} q^n$$

$$(n = 0, 1, 2, \cdots), \tag{17}$$

亦即按最速下降法所得到的逼近序列$\{x_n\}$连同它的各阶导数一致收敛于解 x^*.

更进一步, 如果利用关于用多项式逼近解析函数的 Бернштейн 定理 (Натансон-I, 第 228 页), 则还可以得知解 x^* 是解析的.

现在转入到另一种情形的讨论, 即假定方程(1)的系数和右端函数是任意次可微的函数. 为具体起见, 譬如说假定 a 和 b 具有直到$(\nu+1)$阶的导数, c 和 φ 具有直到 ν 阶的导数, 并且 a 和 b 的$(\nu+1)$阶导数及 c 和 φ 的 ν 阶导数满足指数为 $\alpha>0$ 的 Lipschitz 条件. 由 Jackson 定理(见 Харрик[1]), 可求得次数不超过 n 次的多项式 a_n, b_n, c_n 和 φ_n 使得

$$|a(s, t) - a_n(s, t)| \leqslant \frac{K}{n^{\nu+1+\alpha}},$$

$$\left|\frac{\partial}{\partial s}[a(s, t) - a_n(s, t)]\right| \leqslant \frac{K}{n^{\nu+\alpha}},$$

$$\left|\frac{\partial}{\partial t}[a(s, t) - a_n(s, t)]\right| \leqslant \frac{K}{n^{\nu+\alpha}},$$

$$|b(s,t)-b_n(s,t)|\leqslant\frac{K}{n^{\nu+1+\alpha}},$$

$$\left|\frac{\partial}{\partial s}[b(s,t)-b_n(s,t)]\right|\leqslant\frac{K}{n^{\nu+\alpha}},$$

$$\left|\frac{\partial}{\partial t}[b(s,t)-b_n(s,t)]\right|\leqslant\frac{K}{n^{\nu+\alpha}},$$

$$|c(s,t)-c_n(s,t)|\leqslant\frac{K}{n^{\nu+\alpha}},$$

$$|\varphi(s,t)-\varphi_n(s,t)|\leqslant\frac{K}{n^{\nu+\alpha}}.$$

以 L_n 表示在 L 的表达式中分别用 a_n, b_n, c_n 代替其中的a, b, c 后得到的算子. 易见, 如果 n 充分大, 我们对算子 L_n 应用最速下降法, 并且可以认为量 α 和 β 也适用于算子 L_n. 用 $x^{(n)}$表示方程

$$L_n x=\varphi_n,\qquad\qquad x|_S=0\tag{18}$$

的解, 并来估计这两个解之差

$$\begin{aligned}\|Lx^{(n)}-L_nx^{(n)}\|_{L^2}&\leqslant\left\|\frac{\partial}{\partial s}(a-a_n)\frac{\partial x^{(n)}}{\partial s}\right\|\\&+\left\|(a-a_n)\frac{\partial^2x^{(n)}}{\partial s^2}\right\|+\left\|\frac{\partial}{\partial t}(b-b_n)\frac{\partial x^{(n)}}{\partial t}\right\|\\&+\left\|(b-b_n)\frac{\partial^2x^{(n)}}{\partial t^2}\right\|+\|(c-c_n)x^{(n)}\|\\&\leqslant\frac{K_1}{n^{\nu+\alpha}}\left[\left\|\frac{\partial x^{(n)}}{\partial s}\right\|_{L^2}+\left\|\frac{\partial^2x^{(n)}}{\partial s^2}\right\|_{L^2}+\left\|\frac{\partial x^{(n)}}{\partial t}\right\|_{L^2}\right.\\&\left.+\left\|\frac{\partial^2x^{(n)}}{\partial t^2}\right\|_{L^2}+\|x^{(n)}\|_{L^2}\right]\leqslant\frac{K_2}{n^{\nu+\alpha}}\|x^{(n)}\|_{\mathring{W}_2^{(2)}}.\end{aligned}$$

其次

$$\begin{aligned}\|L(x^{(n)}-x^{(m)})\|_{L^2}&\leqslant\|(L-L_n)x^{(n)}\|+\|(L-L_m)x^{(m)}\|\\&+\|\varphi_n-\varphi_m\|\leqslant\frac{K_2}{n^{\nu+\alpha}}\|x^{(n)}\|_{\mathring{W}_2^{(2)}}+\frac{K_2}{m^{\nu+\alpha}}\|x^{(m)}\|_{\mathring{W}_2^{(2)}}\\&+\frac{K}{n^{\nu+\alpha}}+\frac{K}{m^{\nu+\alpha}}.\end{aligned}$$

根据 Бернштейн-Ладыженская 不等式(见 Ладыженская, 第二章, § 6), 得到

$$\|x^{(n)}-x^{(m)}\|_{\mathring{W}_2^{(2)}}\leqslant K_3\|L(x^{(n)}-x^{(m)})\|_{L^2}$$
$$\leqslant K_4\left[\frac{1}{n^{\nu+\alpha}}\|x^{(n)}\|_{\mathring{W}_2^{(2)}}+\frac{1}{m^{\nu+\alpha}}\|x^{(m)}\|_{\mathring{W}_2^{(2)}}+\frac{1}{n^{\nu+\alpha}}+\frac{1}{m^{\nu+\alpha}}\right]. \tag{19}$$

对 $x^{(n_0)}$ 和 $x^{(m)}$(这里的 n_0 是一个充分大的固定的数) 应用不等式(19)并在其右端用较大的和 $\|x^{(n_0)}\|+\|x^{(n_0)}-x^{(m)}\|$ 代替 $\|x^{(m)}\|$, 即证得量 $\|x^{(n_0)}-x^{(m)}\|_{\mathring{W}_2^{(2)}}$ 的有界性(不依赖于 m) 因而 $\|x^{(m)}\|$ 也有界. 在(19)中令 $n<m$, 利用已经证明了的有界性, 可将这个不等式写成

$$\|x^{(n)}-x^{(m)}\|_{\mathring{W}_2^{(2)}}\leqslant\frac{K_5}{n^{\nu+\alpha}}. \tag{20}$$

取 $h>0$(这个 h 的选择在下面还要讨论)并考察自然数 $n_k=[h^k]$ 的序列 $\{n_k\}$, 其中 $[h^k]$ 乃表示数 h^k $(k=1, 2, \cdots)$ 的整数部分. 对每个函数 $x^{(n_k)}$ 按最速下降法构造 k 阶逼近 $x_k^{(n_k)}$(初值取为零元素). 正如上面已经讲到过的, $x_k^{(n_k)}$ 是次数不超过 $(n_k+2)k$ 次的多项式. 我们来估计两个这样的多项式之差. 首先,

$$\|x_k^{(n_k)}-x_{k-1}^{(n_{k-1})}\|\leqslant\|x_k^{(n_k)}-x^{(n_k)}\|+\|x_{k-1}^{(n_{k-1})}-x^{(n_{k-1})}\|$$
$$+\|x^{(n_k)}-x^{(n_{k-1})}\|.$$

由(20)知

$$\|x^{(n_k)}-x^{(n_{k-1})}\|_{\mathring{W}_2^{(2)}}\leqslant\frac{K_5}{n_{k-1}^{\nu+\alpha}}\leqslant\frac{K_6}{n_k^{\nu+\alpha}}. \tag{21}$$

其次, 由估计式(17)

$$\|x^{(n_k)}-x_k^{(n_k)}\|_{C^{(2)}}\leqslant A_5[(n_k+2)k]^4q^k \quad (k=1, 2, \cdots). \tag{22}$$

因此

$$\|x_k^{(n_k)}-x_{k-1}^{(n_{k-1})}\|_{\mathring{W}_2^{(2)}}\leqslant K_7[(n_k+2)k]^4q^k+\frac{K_6}{n_k^{\nu+\alpha}}.$$

因在范数符号内的表达式是一个次数不超过 $(n_k+2)k$ 的多项式. 所以由引理及不等式(11)可得

$$\|x_k^{(n_k)}-x_{k-1}^{(n_{k-1})}\|_{C^{(p)}}$$
$$\leqslant K_8[(n_k+2)k]^{2p-2}\left\{K_7[(n_k+2)k]^4q^k+\frac{K_6}{n_k^{\nu+\alpha}}\right\}=\varepsilon_k.$$

由于 $n_k\leqslant h^k$, 易见, 若 $\nu+\alpha>2p-2$, 而选择 h 使得有 $h^{2p+2}q<1$, 则级数 $\sum_{k=1}^{\infty}\varepsilon_k$ 收敛. 因此序列 $\{x_k^{(n_k)}\}$ 在空间 $\mathbf{C}^{(p)}$ 中收敛, 并且这个序列的极限是 $\mathbf{C}^{(p)}$ 中的元素. 即 p 次连续可微函数

$$x(s,t)=\lim_{k\to\infty}x_k^{(n_k)}(s,t).$$

如果 $p\geqslant 2$, 利用同一个估计式, 则不难验证 x 满足方程(1), 亦即它是边值问题的古典解. 因此, 如果 $\nu+\alpha>2$, 特别当 $\nu=2$, 即若系数 a 和 b 存在三阶导数, c 和 φ 存在二阶导数, 而且这些导数还满足指数为某 $\alpha>0$ 的 Lipschitz 条件, 则存在古典解.

利用对内子域的估计作类似的讨论我们可得当 $\nu=1$ 时解在内子域中二次连续可微.

最后指出, 这个方法还可用来研究更复杂的问题. 对于 D 不是圆形区域的情形, 如果可经光滑的变换使它变成圆域, 则仍可应用上面的结果. 只要稍加改变, 这个方法还可以应用到具有更多变元的高阶方程上去.

§4. 可微凸泛函的极小化

4.1. 我们在这里来研究最速下降法对于可微泛函极小化的应用. 定义在 Banach 空间 $\mathbf{X}$ 中的泛函 Φ 称作在点 $x\in\mathbf{X}$ 处是可微的, 如果可以找到线性泛函 f 使得对于 $h\in\mathbf{X}$ 成立如下等式

$$\Phi(x+h)=\Phi(x)+f(h)+o(h),$$

其中 $o(h)/\|h\|$ 当 $\|h\|\to 0$ 时趋于零. 这里的泛函 f 就称作泛函 Φ 在点 x 处的 (Fréchet) 导数并记为 $\Phi'(x)$ (关于 Fréchet 微分见第十七章).

设 $\mathbf{X}$ 是 Banach 空间. 从极值问题这个角度考虑, 将 $\mathbf{X}$ 假定为实的 Banach 空间是较为方便的. 这一点是与上一节的情况不同.

我们来考察在 $\mathbf{X}$ 中定义的而且是可微的泛函 Φ, 并说明如何确定这个泛函在某个点 $x\in\mathbf{X}$ 处的最速下降方向. 因为 Φ 是可微的, 所以这个泛函在点 x 处沿方向 z 的导数可以写成

$$\frac{\partial\Phi}{\partial z}(x)=\frac{1}{\|z\|}\lim_{\alpha\to 0+}\frac{1}{\alpha}(\Phi(x+\alpha z)-\Phi(x))$$

$$=\frac{1}{\|z\|}\Phi'(x)(z)=\Phi'(x)\left(\frac{z}{\|z\|}\right).$$

由此可知: 其范数值等于 1 而且其方向取泛函 Φ 在点 x 处的最速下降方向的向量 y 与线性泛函 $\Phi'(x)$ 在单位球面上的极小点相重合, 亦即应由条件

$$\Phi'(x)(y)=\inf_{\|z\|=1}\Phi'(x)(z) \tag{1}$$

来求这样的向量 y. 因为线性泛函不一定在单位球面上取得极小值, 所以最速下降方向可能不存在. 但是, 不难指出使得线性泛函 $\Phi'(x)$ 在单位球面上取得极小值的充分条件. 例如, 当 $\mathbf{X}$ 是共轭于某个 Banach 空间 $\mathbf{Y}$ 而泛函 $\Phi'(x)(x\in\mathbf{X})$ 属于 $\mathbf{Y}$ 时就可保证线性泛函 $\Phi'(x)$ 在单位球面上取得极小值(确切地讲, 这时是取 $\mathbf{Y}$ 的象为 $\mathbf{Y}$ 的正则嵌入, 即 $\mathbf{X}^*=\mathbf{Y}^{**}$). 事实上, 单位球体在 $\mathbf{X}=\mathbf{Y}^*$ 中按(*)-弱拓扑意义下是紧的, 故 $\Phi'(x)$ 在这个球体上达到极小, 从而也就在单位球面上达到极小. 特别, 如果空间 $\mathbf{X}$ 是自反空间, 则最速下降方向显然存在.

下面，我们不限于只考察自反空间这种特殊情形，而是认为空间 X 和泛函 Φ 是使得最速下降方向存在的那种空间和泛函.

如果空间 X 是严格凸的*)，则从(1)可推知最速下降方向是唯一的．当然，一般情况下并非如此.

公式(1)表明，最速下降方向实质上不仅依赖于泛函而且依赖于赋范空间．如果采用别的范数，尽管是与最初的范数等价，则最速下降方向可能会改变．由(1)和等式

$$\inf_{\|z\|=1}\Phi'(x)(z) = -\sup_{\|z\|=1}(-\Phi'(x)(z)) = -\|\Phi'(x)\|$$

推知，方向 $y(\|y\|=1)$是最速下降方向的充分必要条件为

$$\Phi'(x)(y) = -\|\Phi'(x)\|. \tag{2}$$

对于任何 $z\in \mathbf{X}$，成立如下的不等式

$$-\Phi'(x)(z) \leqslant \|\Phi'(x)\|\|z\|.$$

所谓最速下降方向可以这样来刻画：在这个方向上使上述不等式变成了等式．这一点就使得我们可以在某些情况下通过简单的计算找出最速下降方向．试看以下几例:

1) 设 X 是 Hilbert 空间．等式 $-\Phi'(x)(z) = \|\Phi'(x)\|\|z\|$(或者，同样地，等式 $(-\Phi'(x), z) = \|-\Phi'(x)\|\|z\|$)表明向量 $-\Phi'(x)$ 和 z 使得 Cauchy-Буняковский 不等式成为等式．这就说明这两个向量是成正比的，即：存在 $\lambda>0$ 使得 $-\Phi'(x)=\lambda z$．因此，向量 $-\Phi'(x)$就给出了最速下降方向.

注．如果 X 是 Hilbert 空间，则 $\Phi'(x)$可看成同一个空间 X 中的元素，这个元素通常叫作 Φ 在点 x 处的梯度．因此，在这种情

*) Banach 空间X称作严格凸的，如果三角不等式中的等式只对于正的“成比例”元素成立．换言之，从关系式 $\|x+y\|=\|x\|+\|y\|$ 推知对某个 $\lambda>0$ 有 $x=\lambda y$．如果X是严格凸的，$x, y\in$X 并且 $\|x\|=\|y\|=\frac{1}{2}\|x+y\|=1$，则 $x=y$．$L^p(a,b)$ $(1<p<+\infty)$就是严格凸空间一例.

况下，最速下降方向称作反梯度方向.

2) 设 $\mathbf{X}=\mathbf{L}^p(a,b)(1<p<+\infty)$. 分析一下 Hölder 不等式中等号成立的条件，则不难知道 z 在点 x 处的最速下降方向由公式

$$z(t)=-[\mathrm{sign}\,y(t)]|y(t)|^{q-1}$$

给出，其中 $1/p+1/q=1$，y 是空间 $\mathbf{L}^q(a,b)$ 中对所有 $u\in\mathbf{L}^p(a,b)$ 使得等式 $\Phi'(x)(u)=\int_a^b y(t)u(t)\,dt$ 成立的元素. 正象上一种情况一样，在这里，最速下降方向也是唯一的.

3) 设 $\mathbf{X}=\mathbf{L}^\infty(a,b)$ 而泛函 $\Phi'(x)\,(x\in\mathbf{L}^\infty(a,b))$ 属于 $\mathbf{L}^1(a,b)$（也就是说 $\Phi'(x)(u)=\int_a^b y(t)u(t)\,dt$），其中 $y\in\mathbf{L}^1(a,b)$. 在这种情况下，一般来说，最速下降方向是不唯一的. 任何使得 $\mathrm{vrai\,sup}\,|z(t)|=1$ 而且 $z(t)=-\mathrm{sign}\,y(t)$（假定 $y(t)\neq 0$）的可测函数 z 均给出最速下降方向（如果集合 $\{t:\ y(t)=0\}$ 具有零测度，则最速下降方向是唯一的）.

4.2. 在对最速下降法进行研究之前，我们先引入几个定义.

点 $x\in\mathbf{X}$ 称作泛函 Φ 的局部极小点，是指成立着关系 $\Phi(z)\geqslant\Phi(x)$，其中 z 是以 x 为中心的某个球体中的任意元素. 如果对于所有的 $z\in\mathbf{X}$ 均有 $\Phi(z)\geqslant\Phi(x)$，则 x 叫作泛函 Φ 的全局极小点. 点 x 称作可微泛函 Φ 的驻点，是指 $\Phi'(x)=0$.

在研究极小值时，我们在 1.1 中所考察过的实变量的泛函 $\varphi(\alpha;x,z)=\Phi(x+\alpha z)$ 有重要的作用.

不难证明：如果 Φ 是可微泛函，则 $\varphi(\alpha;x,z)$ 对任何 x 和 z 是 α 的可微函数并且 $\varphi'(\alpha;\ x,z)=\Phi'(x+\alpha z)(z)$. 利用这一事实我们来证明：局部极小点是驻点. 事实上，如果在点 x 处达到局部极小，则对任意的 $z\in\mathbf{X}$，函数 $\varphi(\alpha;x,z)$ 在 $\alpha=0$ 的实直线上取得局部极小，因此 $\varphi'(0;x,z)=\Phi'(x)(z)=0$，这即表明 x 是驻点.

逆命题自然不真．但如果Φ是凸泛函，则驻点就是全局极小点．

给定在线性空间 $\mathbf{X}$ 上的泛函Φ称作凸泛函，是指对于任意的 $z_1, z_2 \in \mathbf{X}$ 和实数 $t_1, t_2 \geqslant 0 (t_1 + t_2 = 1)$ 均成立如下的不等式

$$\Phi(t_1 x_1 + t_2 x_2) \leqslant t_1 \Phi(x_1) + t_2 \Phi(x_2).$$

如果Φ是凸泛函，则对任意的 x 和 z 来说 $\varphi(\alpha; x, z)$ 是实直线上的凸函数，亦即

$$\varphi(t_1\alpha_1 + t_2\alpha_2; x, z) \leqslant t_1 \varphi(\alpha_1; x, z) + t_2 \varphi(\alpha_2; x, z);$$

这里的 $t_1, t_2, \alpha_1, \alpha_2$ 均是实数，而且 $t_1, t_2 \geqslant 0, t_1 + t_2 = 1$．从 $\varphi(\alpha; x, z)$ 的凸性可知这个函数的导数不下降．

引理 1． 设Φ是凸可微泛函．那么，Φ的每个驻点均是该泛函的全局极小点．

证． 设 x 是驻点，z 是 $\mathbf{X}$ 中的任意一个元素．因为函数 $\varphi(\alpha; x, z)$ 的导数不下降，所以利用 Lagrange 公式就有

$$\begin{aligned}\Phi(x+z) - \Phi(x) &= \varphi(1; x, z) - \varphi(0; x, z) \\ &= \varphi'(\theta; x, z) \geqslant \varphi'(0; x, z) = \Phi'(x)(z) = 0\end{aligned}$$

(其中 $0 < \theta < 1$)．这表明对所有的 $z \in \mathbf{X}$ 均成立关系式 $\Phi(x+z) \geqslant \Phi(x)$，这正是所要证明的．

下面将谈到最速下降法的收敛性问题，即这个方法收敛于泛函Φ的驻点的问题(当Φ是凸泛函时就是收敛于它的全局极小点的问题)．

关于收敛性的基本结果只有当泛函Φ的导数在某个球体中满足 Lipschitz 条件，即存在数 L 使得对这个球体中的 x 和 z 有如下不等式

$$\|\Phi'(x) - \Phi'(z)\| \leqslant L \|x - z\| \tag{3}$$

时才成立．这时要用到下面的引理．

引理 2． 设导数 Φ' 满足 Lipschitz 条件(3)，其中 x 和 z 属于

以零点为中心以 $a+b$ 为半径的球体 $B_{a+b}(0)$. 这时，如果 $\|x\|\leqslant a$，则

$$\Phi(x+z)\leqslant\Phi(z)+\Phi'(x)(z)+\frac{L}{2}\|z\|^2.$$

证. 利用函数 $\varphi(\alpha;x,\ z)$ 的可微性和这个函数的导数的表达式，就有

$$\begin{aligned}\Phi(x+z)&=\varphi(1;x,z)=\varphi(0;x,z)+\int_0^1\varphi'(\alpha;x,z)d\alpha\\&=\Phi(x)+\int_0^1\Phi'(x+\alpha z)(z)d\alpha\\&=\Phi(x)+\Phi'(x)(z)+\int_0^1(\Phi'(x+\alpha z)-\Phi'(x))(z)d\alpha\\&\leqslant\Phi(x)+\Phi'(x)(z)+\int_0^1\|\Phi'(x+\alpha z)-\Phi'(x)\|\|z\|d\alpha\\&\leqslant\Phi(x)+\Phi'(x)(z)+L\|z\|^2\int_0^1\alpha d\alpha\\&=\Phi(x)+\Phi'(x)(z)+(L/2)\|z\|^2.\end{aligned}$$

引理得证.

4.3. 我们应用最速下降法于泛函 Φ 的极值问题. 换言之，我们来研究序列 $x_0,x_1,\cdots,x_n,\cdots$ 的情况，其中

$$x_n=x_{n-1}+\varepsilon_n z_n\qquad(n=1,2,\cdots),\tag{4}$$

z_n 是泛函 Φ 在点 x_{n-1} 处的最速下降方向（如果有多个最速下降方向，则从中任取一个作为 z_n），而 ε_n 从条件

$$\Phi(x_{n-1}+\varepsilon_n z_n)=\min_{\alpha\geqslant0}\Phi(x_{n-1}+\alpha z_n)$$

求得. 因此，下降值按 1.1 中所指出的第二个方法选取. 当然，如果 Φ 是严格凸泛函，这两种方法是一样的（Φ 的严格凸性乃表明：对于 $x_1,x_2\in\mathbf{X}$，$x_1\neq x_2$；$t_1,t_2>0$，$t_1+t_2=1$ 成立严格的不等关系 $\Phi(t_1x_1+t_2x_2)<t_1\Phi(x_1)+t_2\Phi(x_2)$；从 Φ 的严格凸性推知函数

$\varphi(\alpha; x, z)$的导数 $\varphi'(\alpha; x, z)$是严格上升的，因而 $\varphi'(\alpha; x, z)$只能在一点变成零，由此就推出了这两个方法的等价性）.

下面我们不再特别说明而假定：初始点 $x_0 \in \mathbf{X}$ 使得 Lebesgue 集合 $\Omega_0 = \{x \in \mathbf{X}: \Phi(x) \leqslant \Phi(x_0)\}$有界. 还要指出的一点是：$x_n \in \Omega_0$ $(n=0, 1, 2, \cdots)$. 同样，还将假定对所有的 n 均有 $\|z_n\| = 1$.

定理 1. 设泛函Φ的导数在以零点为中心而以 $R'(R' > R = \sup\limits_{x \in \Omega_0} \|x\|)$ 为半径的球体中满足 Lipschitz 条件(3). 那么，按公式(4)构成的序列$\{x_n\}$使得 $\Phi'(x_n) \to 0$.

证. 设 $0 < \alpha \leqslant R' - R$. 按照定义，下降值 ε_n 满足 $\Phi(x_{n-1} + \alpha z_n) \geqslant \Phi(x_n)$. 利用引理 2，我们就得到

$$\begin{aligned}\Phi(x_n) &\leqslant \Phi(x_{n-1} + \alpha z_n) \\ &\leqslant \Phi(x_{n-1}) - \alpha \Phi'(x_{n-1})(z_n) + \frac{1}{2} L \alpha^2\end{aligned}$$

（其中 L 是一常数，由 Lipschitz 条件来确定）. 由这个不等式以及关系式(2)推知

$$\|\Phi'(x_n)\| \leqslant \frac{\Phi(x_{n-1}) - \Phi(x_n)}{\alpha} + \frac{1}{2} L \alpha .$$

设 ε 是任意的正数而 $\alpha < \min(\varepsilon, R' - R)$. 从 Ω_0 的有界性和 Lipschitz 条件推知泛函Φ下有界. 因为序列 $\Phi(x_n)$下降（按其构造）并且有界，所以这个序列收敛，因此对充分大的 n 将有$\frac{1}{\alpha}(\Phi(x_{n-1}) - \Phi(x_n)) < \varepsilon$. 对使这个不等式成立的 n 来说，就有

$$\|\Phi'(x_n)\| \leqslant \varepsilon \left(1 + \frac{1}{2} L\right),$$

这就完成了定理的证明.

推论. 序列(4)的极限点（如果它存在的话）是驻点.

在某些情况下，基于紧性的考虑，可以证明，对于其导数仅仅连续的泛函来说，其最速下降序列收敛于该泛函的驻点. 我们给

出一个这样的结果(见 Curry[1]).

定理 2. 设 X 是有限维空间而 Φ 是连续可微泛函. 那么, 相应的最速下降序列 $\{x_n\}$ 的极限点是驻点.

证. 首先指出: 从空间 X 是有限维空间以及 Ω_0 是有界集这两个事实即可推知序列 $\{x_n\}$ 的极限点存在. 设 y 是这样的一个极限点, 即 $y=\lim x_{n_k}$. 不失一般性, 可以认为存在着极限 $\lim z_{n_k+1}=\tilde{y}$. 对正的 α, 设

$$w_{n_k}(\alpha)=\frac{1}{\alpha}(\Phi(x_{n_k}+\alpha z_{n_k+1})-\Phi(x_{n_k}))-\Phi'(x_{n_k})(z_{n_k+1}). \tag{5}$$

从定理的条件推知存在 $\lim\limits_{k\to\infty} w_{n_k}(\alpha)=w(\alpha)$, 并且

$$w(\alpha)=\frac{1}{\alpha}(\Phi(y+\alpha\tilde{y})-\Phi(y))-\Phi'(y)(\tilde{y}).$$

因为 Φ 在点 y 处可微, 所以 $\lim\limits_{\alpha\to 0} w(\alpha)=0$. 由于下降值的选择,

$$\Phi(x_{n_k+1})=\Phi(x_{n_k}+\varepsilon_{n_k+1}z_{n_k+1})\leqslant\Phi(x_{n_k}+\alpha z_{n_k+1}),$$

因而再由(5), 即得

$$\Phi(x_{n_k+1})\leqslant\Phi(x_{n_k})+\alpha\Phi'(x_{n_k})(z_{n_k+1})+\alpha w_{n_k}(\alpha).$$

因为 $\alpha>0$, 所以

$$-\Phi'(x_{n_k})(z_{n_k+1})\leqslant\frac{1}{\alpha}(\Phi(x_{n_k})-\Phi(x_{n_k+1}))+w_{n_k}(\alpha). \tag{6}$$

在(6)式中令 $k\to\infty$ 取极限, 考虑到 $\lim(\Phi(x_{n_k})-\Phi(x_{n_k+1}))=0$ 并利用(2), 我们就得到

$$\|\Phi'(y)\|=-\Phi'(y)\tilde{y}\leqslant w(\alpha).$$

由于 $\lim\limits_{\alpha\to 0} w(\alpha)=0$, 从而定理得证.

4.4. 如果 Φ 是凸泛函, 则可进一步指出最速下降法的收敛速度.

引理 3. 设Φ是在 $\mathbf{X}$ 上下有界的凸泛函并且 $Q=\inf\limits_{x\in\mathbf{X}}\Phi(x)$. 那么, 存在数 $c>0$ 使得

$$\Phi(x_n)-Q\leqslant c\|\Phi'(x_n)\|.$$

证. 因为集合 $\Omega_0=\{z\in\mathbf{X}: \Phi(z)\leqslant\Phi(x_0)\}$ 按照假定是有界的, 所以集合 $\Omega_0-\Omega_0=\{z: z_1=z_1-z_2z_1, z_2\in\Omega_0\}$ 同样是有界的. 我们来证明: c 可以取为以零点为中心且包含 $\Omega_0-\Omega_0$ 的球B的半径. 考虑到函数 $\varphi(\alpha; x_n, z)$的导数 $\varphi'(\alpha; x_n, z)$不上升(由于Φ是凸泛函)并且利用 Lagrange 公式, 我们就得到

$$\begin{aligned}\Phi(x_n+z)-\Phi(x_n)&=\varphi(1; x_n, z)-\varphi(0; x_n, z)\\&=\varphi'(\theta; x_n, z)\geqslant\varphi'(0; x_n, z)=\Phi'(x_n)(z)\end{aligned}$$

(其中 $0<\theta<1$). 由此得到

$$\min_{\|z\|\leqslant c}\Phi(x_n+z)-\Phi(x_n)\geqslant\min_{\|z\|\leqslant c}\Phi'(x_n)(z). \tag{7}$$

因为 $x_n\in\Omega_0$, 所以 $B\supset\Omega_0-\Omega_0\supset\Omega_0-x_n$, 也就是 $x_n+B\supset\Omega_0$. 因此

$$Q=\inf_{x\in\mathbf{X}}\Phi(x)\leqslant\inf_{\|z\|\leqslant c}\Phi(x_n+z)\leqslant\inf_{x\in\Omega_0}\Phi(x).$$

从 Ω_0 的定义可推知 $Q=\inf\limits_{x\in\Omega_0}\Phi(x)$. 此外, 还有

$$\inf_{\|z\|\leqslant c}\Phi'(x_n)(z)=c\inf_{\|z\|\leqslant 1}\Phi'(x_n)(z)=-c\|\Phi'(x_n)\|.$$

从上面得到的关系式及(7)推知有

$$Q-\Phi(x_n)\geqslant-c\|\Phi'(x_n)\|.$$

引理得证.

推论. 如果引理 3 的条件成立并且定理 1 或定理 2 的条件也成立, 则序列$\{x_n\}$是极小化序列.

在导数 $\Phi'(x)$满足 Lipschitz 条件的假定下可以得到序列$\{\Phi(x_n)\}$的收敛速度的估计.

定理 3. 设Φ是在 $\mathbf{X}$ 上下有界的凸泛函并且 $\Phi'(x)$在 $\mathbf{X}$ 上满足 Lipschitz 条件(3). 在这些假定之下, 成立如下估计式

$$\Phi(x_n)-Q=O(1/n),$$

其中 $Q=\inf\limits_{x\in \mathbf{X}}\Phi(x)$.

证. 证明要依据下面的引理.

引理 4. 设数 $\lambda_n>0(n=1,2,\cdots)$具有这样的性质: 对某个 $\mu>0$ 和所有的 n 满足不等式 $\lambda_n-\lambda_{n+1}\geqslant\mu\lambda_n^2$. 那么, 就有 $\lambda_n=O\left(\frac{1}{n}\right)$.

证. 因为序列 λ_n 是下降的, 故

$$\lambda_n-\lambda_{n+1}=\lambda_{n+1}\left(\frac{\lambda_n}{\lambda_{n+1}}-1\right)\geqslant\mu\lambda_{n+1}^2,$$

由此推知 $\frac{\lambda_n}{\lambda_{n+1}}-1\geqslant\mu\lambda_{n+1}$.

令 $\nu_n=\lambda_n n$, 由上一不等式再经简单的变换就得到

$$\frac{\nu_n}{\nu_{n+1}}\geqslant\frac{n}{n+1}\left(1+\mu\frac{\nu_{n+1}}{n+1}\right). \tag{8}$$

如果对某个 n 有 $\nu_{n+1}\geqslant 2/\mu$, 则由(8)推知 $\nu_{n+1}\leqslant\nu_n$. 这样一来就有 $\nu_{n+1}\leqslant\max(2/\mu,\nu_n)$, 由此推知对所有的 n 均有 $\nu_n\leqslant\max(2/\mu,\nu_1)$. 引理证毕.

定理 3 的证明. 令 $\lambda_n=\Phi(x_n)-Q$ 并假定对所有的 n 均有 $\lambda_n>0$. 我们来证明这样的 λ_n 对某个 $\mu>0$ 满足引理 4 的条件. 为此, 我们首先来估计 $\lambda_n-\lambda_{n+1}=\Phi(x_n)-\Phi(x_{n+1})$. 从引理 2 推知, 对所有的实数 α 均有

$$\Phi(x_n+\alpha z_{n+1})\leqslant\Phi(x_n)+\alpha\Phi'(x_n)(z_{n+1})+\frac{\alpha^2}{2}L\|z_{n+1}\|^2.$$

因为 $\|z_{n+1}\|=1$ 而 $\Phi'(x_n)(z_{n+1})=-\|\Phi'(x_n)\|$, 所以

$$\Phi(x_n+\alpha z_{n+1})\leqslant\Phi(x_n)-\alpha\|\Phi'(x_n)\|+\frac{\alpha^2}{2}L. \tag{9}$$

设 $\tilde{\alpha}=\frac{1}{L}\|\Phi'(x_n)\|$((9)式右端的二次三项式在这一点取得极小值). 由(9)式并考虑到 $\tilde{\alpha}$ 的定义, 我们就得到

$$\lambda_n-\lambda_{n+1}=\Phi(x_n)-\Phi(x_{n+1})\geqslant\Phi(x_n)-\Phi(x_n+\tilde{\alpha}z_{n+1})$$

$$\geqslant\tilde{\alpha}\|\Phi'(x_n)\|+\frac{\tilde{\alpha}^2}{2}L=\frac{L}{2}\|\Phi'(x_n)\|^2.$$

另一方面，由引理 3 可知，对某个 $c>0$ 将成立不等式

$$\lambda_n=\Phi(x_n)-Q\leqslant c\|\Phi'(x_n)\| \qquad (n=1,2,\cdots).$$

因此，对所有的 n

$$\lambda_n-\lambda_{n+1}\geqslant\frac{L}{2c^2}\lambda_n^2.$$

考虑到引理 4 的结果就完成了定理的证明.

对于在定理 3 中所考察的广泛的一类泛函来说，其收敛性的估计未必能再得到改善. 但对于"好的"泛函来说，方法的收敛性实际上是更快一些. 定义在 Hilbert 空间 $\mathbf{H}$ 上的二次可微泛函 Φ 称作是强凸的，是指存在正数 M 和 m 使得对所有的 x, $z\in\mathbf{H}$ 均有

$$m\|z\|^2\leqslant(\Phi''(x)(z),z)\leqslant M\|z\|^2. \tag{10}$$

这里的 $\Phi''(x)$ 是泛函 Φ 在点 x 处的二阶导数*). 不难验证：泛函 Φ 的强凸性蕴含着 Φ 的凸性. 不等式(10)实质上表明算子 $\Phi''(x)$ 对所有的 $x\in\mathbf{H}$ 是正定的而且这些算子的界一致地囿于正的常数之间. 特别，在 §1 中所考察过的二次泛函 F 是强凸的（对于所有的 x，均有 $F''(x)=U$).

对于强凸泛函的最速下降法来说，很多结果都与在 §1 中对二次泛函所得到的结果相类似. 可以证明，对任意的 $x_0\in\mathbf{H}$ 集合 $\Omega_0=\{z\in\mathbf{H},\Phi(z)\leqslant\Phi(x_0)\}$ 均有界，而且泛函 Φ 在这个集合上取得极小值（由于 x_0 的任意性，实际上是在整个空间上取得极小值），在这种情况下，极小点 x^* 是唯一的. 正如对二次泛函一样，成立着如下的收敛性估计

---

*) $\Phi''(x)$ 是从 $\mathbf{X}=\mathbf{H}$ 到 $\mathbf{X}^*$ 的线性算子，它由关系式 $\Phi'(x+h)=\Phi'(x)+\Phi''(x)(h)+o(h)$ 确定，其中 $\frac{o(h)}{\|h\|}\xrightarrow[h\to0]{}0$. 详见第十七章.

$$\|x_n - x^*\| \leqslant C\left(\frac{M-m}{M+m}\right)^n,$$

其中 C 是依赖于初始点 x_0 和泛函 Φ 的一个常数.

4.5. 最速下降这一思想应用于条件极值问题的求解是最富有成效的. 应用最速下降思想于条件极值问题的解法之一即归之于所谓的条件梯度法.

设 Ω 是 Banach 空间 $\mathbf{X}$ 中的凸弱紧集合. 假定集合 Ω 的结构极为简单,即线性泛函在 Ω 上的极小化问题的解是已知的. 其次,假定 Φ 是定义在包含 Ω 的某个开域上的可微(非线性)泛函. 现在的问题是要求泛函 Φ 在 Ω 上的极小值. 条件梯度法使得我们可以作出一个序列 $x_0, x_1, x_2, \cdots, x_n, \cdots$, 这个序列在很多情形下都是极小化序列. 初始值 x_0 可取 Ω 中任意元素. 如果 x_n 已经得到, 则 x_{n+1} 可按如下公式来确定

$$x_{n+1} = x_n + \varepsilon_n(z_n - x_n),$$

其中 $z_n \in \Omega$ 由条件 $\Phi'(x_n)(z_n) = \min\limits_{z \in \Omega} \Phi'(x_n)(z)$ 来选取, 下降值 ε_n 则由等式 $\Phi(x_{n+1}) = \min\limits_{\alpha \in [0,1]} \Phi(x_n + \alpha(z_n - x_n))$ 来求得. (确定下降值还有其他可能的方法,但对此我们就不再考虑了). 因此, 如果在最速下降法中按使 $\Phi'(x)$ 在单位球面上(球面的方向)达到极小的原则求下降的方向, 而按使泛函在射线方向上达到极小的原则求下降值, 则在条件梯度法中就是按使 $\Phi'(x)$ 在集合 Ω 上达到极小的原则求下降的方向,而按使泛函 Φ 在区间 $[0, 1]$ 上达到极小的原则求下降值. 粗略地讲, 条件梯度法乃是在以上限制之下求极值的最速下降法的变型.

对于条件梯度法, 类似于定理 1, 2, 3 的事实也成立. 这时, 点 $y \in \Omega$ 称作泛函 Φ 在集合 Ω 上的驻点, 是指 $\Phi'(y)(y) = \min\limits_{z \in \Omega} \Phi'(y)(z)$. 容易证明: Φ 在 Ω 上的局部极小点是驻点; 如果 Φ 是凸泛函, 则

Φ 在 Ω 上的驻点亦是 Φ 的全局极小点. 如果 $\mathbf{X}$ 是 Hilbert 空间, Φ 是凸泛函, $\inf\limits_{x\in\Omega}\|\Phi'(x)\|>0$, Ω 是强凸集(亦即存在 $\nu>0$ 使得对于每个 $z,x\in\Omega$ 均有 $\frac{x+z}{2}+\nu u\in\Omega$, 这里 $\|u\|\leqslant\|x-z\|^2$), 则按条件梯度法所构成的序列以几何级数的速度收敛于 Φ 在 Ω 上的唯一极小点. 关于条件梯度法的详细叙述可参见 Демьянов 和 Рубинов.

§5. 有限维空间凸泛函的极小化

5.1. 我们在 §4 中已经考察过连续可微泛函的极值问题中的最速下降法. 下面来研究有限维欧氏空间 $\mathbf{R}^n$ 上凸泛函的极值问题.

设 Φ 是任意的凸泛函*). 众所周知(例如, 参见 Рокафеллар), $\mathbf{R}^n$ 上的凸泛函是连续的.

向量 $v\in\mathbf{R}^n$ 称为泛函 Φ 在点 x 处的次梯度, 乃指对所有的 $z\in\mathbf{R}^n$ 均有

$$\Phi(z)\geqslant\Phi(x)+(v,z-x). \tag{1}$$

次梯度的集合是非空的、凸的、闭的和有界的. 这个集合称作泛函 Φ 在点 x 处的次微分并以 $\partial\Phi(x)$ 来表示. 次微分乃是一种多值映射, 对此将在 XVI, §5 中详细论及. 我们在这里只指出一个事实, 即映射 $\partial\Phi(x)$ 是上半连续的: 如果 $x_k\to x$, $v_k\to v$, 并且 $v_k\in\partial\Phi(x_k)$, 则 $v\in\partial\Phi(x)$.

凸泛函 Φ 是按方向可微的, 即对于任意的 $x\in\mathbf{R}^n$ 和 $y\in\mathbf{R}^n$, 存在

$$\frac{1}{\|y\|}\lim_{\alpha\to+0}\frac{1}{\alpha}[\Phi(x+\alpha y)-\Phi(x)]\equiv\frac{\partial\Phi}{\partial y}(x),$$

并且

---

*) 凸泛函的定义见 4.2 所述.

$$\|y\|\frac{\partial \Phi}{\partial y}(x)=\max_{v\in\partial\Phi(x)}(v,y). \tag{2}$$

如果$\|y\|=1$并且

$$\frac{\partial \Phi}{\partial y(x)}(x)=\inf_{\|y\|=1}\frac{\partial \Phi}{\partial y}(x),$$

则称方向$y(x)$是泛函Φ在点x处的最速下降方向.

从(2)可以看出：如果$\inf\limits_{\|y\|=1}\frac{\partial \Phi}{\partial y}(x)<0$，则最速下降方向存在、唯一而且可以由公式

$$y(x)=-\frac{z(x)}{\|z(x)\|} \tag{3}$$

求出，其中$\|z(x)\|=\min\limits_{z\in\partial\Phi(x)}\|z\|\equiv\rho(x)$.

引理 1. 凸泛函Φ在点x处达到它在$\mathbf{R}^n$上的最小值的充分必要条件是

$$\inf_{\|y\|=1}\frac{\partial \Phi}{\partial y}(x)\geqslant 0. \tag{4}$$

证. 必要性显然，我们只证充分性．也就是说要来证明当(4)式被满足时，相应的x是极小点．用反证法．设存在点z使得$\Phi(z)<\Phi(x)$．我们来考察方向$y=\frac{z-x}{\|z-x\|}$．按照凸泛函的定义，对于小的α将有

$$\Phi(x+\alpha y)=\Phi\left(\left(1-\frac{\alpha}{\|z-x\|}\right)x+\frac{\alpha}{\|z-x\|}z\right)$$

$$\leqslant\left(1-\frac{\alpha}{\|z-x\|}\right)\Phi(x)+\frac{\alpha}{\|z-x\|}\Phi(z),$$

因此

$$\frac{\partial \Phi}{\partial y}(x)\equiv\lim_{\alpha\to+0}\frac{1}{\alpha}[\Phi(x+\alpha y)-\Phi(x)]$$

$$\leqslant\frac{1}{\|z-x\|}(\Phi(z)-\Phi(x))<0,$$

这与(4)相矛盾.

从几何上讲，条件(4)等价于

$$\mathbf{0}\in\partial\Phi(x). \tag{5}$$

条件(5)是连续可微泛函极值的必要条件 $\Phi'_x=\mathbf{0}$ 的推广.

我们举一例来说明. 设

$$\Phi(x)=\max_{z\in G}F(x,z), \tag{6}$$

其中 G 是紧集合，泛函 $F(x,z)$ 连同 $F'_x(x,z)$ 在乘积空间 $\mathbf{R}^n\times G$ 上是连续的而且对于每个确定的 $z\in G$ 来说 $F(x,z)$ 和 $F'_x(x,z)$ 是 x 的凸泛函. 这样的泛函 Φ 是凸泛函，并且

$$\partial\Phi(x)=\mathrm{co}\{F'_x(x,z):\ z\in R(x)\},$$

其中

$$R(x)=\{z\in G:\ F(x,z)=\Phi(x)\},$$

$\mathrm{co}A$ 表示集合 A 的凸包.

注. 泛函(6)即使当 F 关于 x 不是凸的，一般来说，也是按方向可微的. 此外，尽管 $\partial\Phi(x)$ 已不再是次微分，公式(2)依然成立.

设 $\varepsilon>0$. 如果对所有的 $z\in\mathbf{R}^n$ 均有

$$\Phi(z)\geqslant\Phi(x)+(v,z-x)-\varepsilon, \tag{7}$$

则称向量 v 是泛函 Φ 在点 x 处的ε-次梯度.

所有 ε-次梯度的集合记为 $\partial_\varepsilon\Phi(x)$. 这个集合是闭的、凸的和有界的. 如果

$$\mathbf{0}\in\partial_\varepsilon\Phi(x), \tag{8}$$

则称 x 为泛函 Φ 的ε-驻点. 由(7)易于看到有

$$0\leqslant\Phi(x)-\min_{z\in\mathbf{R}^n}\Phi(z)\leqslant\varepsilon. \tag{9}$$

如果 $\mathbf{0}\notin\partial_\varepsilon\Phi(x)$，命 $\|z_\varepsilon(x)\|=\min_{z\in\partial_\varepsilon\Phi(x)}\|z\|\equiv\rho_\varepsilon(x)$，则称方向 $y_\varepsilon(x)=-\dfrac{z_\varepsilon(x)}{\|z_\varepsilon(x)\|}$ 为泛函 Φ 在点 x 处的ε-最速下降方向. 映射 $\partial_\varepsilon\Phi(x)$

也是上半连续的.

设 $v\in\partial\Phi(x)$. 从(1)可得到

$$\begin{aligned}\Phi(z)&\geqslant\Phi(x)+(z-x,v)\\&=\Phi(x_0)+(v,z-x_0)+[\Phi(x)-\Phi(x_0)+(v,x_0-x)].\end{aligned}$$

由此可见, 对于充分接近 x_0 的 x 有

$$\partial\Phi(x)\subset\partial_\varepsilon\Phi(x_0). \tag{10}$$

5.2. 很自然地, 我们要考察最速下降法在这种情形中的应用.

任取 $x_0\in\mathbf{R}^n$. 设已经求得 $x_k\in\mathbf{R}^n$. 如果 $\mathbf{0}\in\partial\Phi(x_k)$, 则 x_k 就是极小点, 于是最速下降过程结束. 如果 $\mathbf{0}\notin\partial\Phi(x_k)$, 则可求得 $y_{k+1}=y(x_k)$(见式(3)). y_{k+1} 乃是泛函 Φ 在点 x_k 处的最速下降方向. 我们来考察射线

$$x_{k\alpha}=x_k+\alpha y_{k+1}\qquad(\alpha\geqslant0)$$

并求

$$\min_{\alpha\geqslant0}\Phi(x_{k\alpha})=\Phi(x_{k\alpha_k}).$$

命 $x_{k+1}=x_{k\alpha_k}$, 则易见有 $\Phi(x_{k+1})<\Phi(x_k)$.

但是, 刚才所述的这个方法可能不收敛于极小点("梗阻"效应), 这表明映射 $\partial\Phi(x)$ 的间断性(相应的例子见 Демьянов 和 Малозамов 的书).

5.3. 设 $\varepsilon>0$. 为了求 ε-驻点可以采用下面的方法. 任取 $x_0\in\mathbf{R}^n$. 假定集合 $\Omega_0=\{x:\ \Phi(x)\leqslant\Phi(x_0)\}$ 有界. 设已经求得 $x_k\in\mathbf{R}^n$. 如果 $\mathbf{0}\in\partial_\varepsilon\Phi(x_k)$, 则 x_k 就是 ε-驻点, 从而这一过程结束. 如果 $\mathbf{0}\notin\partial_\varepsilon\Phi(x_k)$, 则我们可求得 $y_{k+1}=y_\varepsilon(x_k)$并来考察射线 $x_{k\alpha}=x_k+\alpha y_{k+1}(\alpha\geqslant0)$. 先求得 $\min\limits_{\alpha\geqslant0}\Phi(x_{k\alpha})=\Phi(x_{k\alpha_k})$, 再命 $x_{k+1}=x_{k\alpha_k}$, 则容易看到有 $\Phi(x_{k+1})<\Phi(x_k)$.

类似地，设我们构造出了序列$\{x_k\}$，如果这个序列包含有限个点，则最后所得到的那个点就是ε-驻点，如果这个序列是无限的，则成立如下的定理．

定理 1． 序列$\{x_k\}$的所有极限点都是泛函Φ的ε-驻点．

证． 设$x_{k_s}\to\tilde{x}$．应当证明$\mathbf{0}\in\partial_\varepsilon\Phi(\tilde{x})$．假定不是这样，即$\mathbf{0}\notin\partial_\varepsilon\Phi(\tilde{x})$．设$\rho_\varepsilon(\tilde{x})=a>0$．从映射$\partial_\varepsilon\Phi(x)$的上半连续性推知存在$\delta>0$使得对于所有的

$$x\in S_\delta(\tilde{x})=\{x:\ \|x-\tilde{x}\|\leqslant\delta\}$$

均有$\rho_\varepsilon(x)\geqslant\dfrac{a}{2}$.

我们知道，对充分大的k_s将有

$$\rho_\varepsilon(x_{k_s})\geqslant\frac{a}{2}. \tag{11}$$

从(10)式可知，存在$\delta_1>0$使得对于所有的$x\in S_{\delta_1}(x_{k_s})$以及充分大的$k_s$均有$\partial\Phi(x)\subset\partial_\varepsilon\Phi(x_{k_s})$．于是，泛函$\Phi$在射线$x_{k_s\tau}$上以按绝对值不小于$a/2$的速度下降而且其间的距离不小于$\delta_1$．因而

$$\Phi(x_{k_s+1})\leqslant\Phi(x_{k_s})-\delta_1\frac{a}{2}.$$

由此得出$\Phi(x_k)\to-\infty$，这与关于集合Ω_0的有界性假定相矛盾．

5．4． 因此，求泛函Φ的极小点的过程可以是这样的：任取$\varepsilon_0>0$, $\rho_0>0$, $x_0\in\mathbf{R}^n$．设已经求得$\varepsilon_k>0$，$\rho_k>0$，$x_k\in\mathbf{R}^n$．如果$\mathbf{0}\in\partial\Phi(x_k)$，则$x_k$就是极小点．不然的话，应用5．3段中所述的方法(取x_k为初始点)，经有限步后我们就可得到x_{k+1}使得$\rho_{\varepsilon_k}(x_{k+1})\leqslant\rho_k$．现在令$\varepsilon_{k+1}=\beta\varepsilon_k$, $\rho_{k+1}=\beta\rho_k$，其中$\beta\in(0,1)$且不依赖于k．因此，我们就构造出了序列$\{x_k\}$．不难证明(见Демьянов和Малоземов的书)下述定理．

定理 2． 序列$\{x_k\}$的所有极限点都是泛函Φ的极小点．

注． 以上所述只是算法的"原则性"格式．在实际应用中，象

求下降方向、求在射线上的极小值等辅助问题的解都是近似得到的.

5.5. 我们现在来考察广义梯度法(见Ермольев和Шор[1]). 设Φ是$\mathbf{R}^n$上的凸泛函并且在$\mathbf{R}^n$上达到它的极小, M是极小点组成的集合, $\rho(x)=\min\limits_{z\in M}\|z-x\|$. 以$M_\varepsilon$表示集合$M$的$\varepsilon$邻域, 即

$$M_\varepsilon=\{x:\ \rho(x)\leqslant\varepsilon\},\qquad \varepsilon>0.$$

设$\partial\Phi(x)$是泛函Φ在点x处的次微分, 即

$$\partial\Phi(x)\equiv\{z\in\mathbf{R}^n:\ \Phi(y)-\Phi(x)\geqslant(z,y-x)\text{对所有的}\ y\in\mathbf{R}^n\}.$$

取$x_0\in\mathbf{R}^n$和正数序列$\{\lambda_k\}$, 使得

$$\lambda_k\to0,\ \sum_{k=0}^{\infty}\lambda_k=\infty.$$

按如下公式构造序列$\{x_k\}$:

$$x_{k+1}=x_k-\lambda_k y_{k+1},\qquad y_{k+1}=z_{k+1}/\|z_{k+1}\|,$$

其中z_{k+1}是$\partial\Phi(x_k)$中任意向量. 如果到某一步有$z_k=\mathbf{0}$, 则$x_k\in M$, 此过程即结束.

我们现在假定$\{x_k\}$是无限序列.

定理 3. 如果集合M有界, 则

$$\rho(x_k)\to0,\ \Phi(x_k)\to\tilde{\Phi}\equiv\min_{x\in\mathbf{R}^n}\Phi(x).$$

证. 设$\Omega(x)=\{y:\ \Phi(y)\leqslant\Phi(x)\}$. 因为$M$有界, 所以$\Omega(x)$对于任意的$x\in\mathbf{R}^n$同样是有界的(见Rokafellar的专著"凸分析"). 命

$$T(x)=\{y:\ \Phi(y)=\Phi(x)\},\ b(x)=\min_{y\in T(x)}\rho(y).$$

不难证明

$$M_{b(x)}\subset\Omega(x).\tag{12}$$

设$x\notin M$. 那么就有$b(x)>0$. 取任意方向$y=\frac{z}{\|z\|}$, $z\in\partial\Phi(x)$. 对

所有的 $v\in\Omega(x)$，均有

$$(y, v-x)\leqslant\frac{1}{\|z\|}(\Phi(v)-\Phi(x))\leqslant 0.$$

特别，对于所有的 $z\in M$，由于(12)，则

$$0\geqslant(y, z+b(x)y-x).$$

由此不难得知泛函 ρ 在方向 $-y$ 上是下降的而且

$$\lim_{k\to\infty}b(x_k)=0. \tag{13}$$

由(13)式可知存在子序列$\{x_{k_s}\}$使得 $b(x_{k_s})\to 0$. 从而就有 $\rho(x_{k_s})\to 0$. 事实上，可以证明整个序列 $\rho(x_k)$收敛于 0，亦即 $\Phi(x_k)\to\tilde{\Phi}$.

注. 当Φ是连续可微泛函时，其次微分由唯一的一个点（它的梯度)组成. 如果按照 1.1 中所述的第二种方法选择下降值，则 5.4 中所说的最速下降法与 § 4 中所考察的最速下降法是一致的. 如果按 1.1 中所述的第三种方法选取下降值，则 5.5 中所述的广义梯度法就变成了最速下降法.

Демьянов 和 Малоземов 的专著详细讨论了最速下降法，关于广义梯度法，可在 Ермольев 和 Шор[1]的著作中找到.

本节所述的方法已被推广到求有界泛函的极值问题上去. 这种推广乃是依据极小值的必要条件(见 Дубовицкий 和Милютин [1]的著作以及 Пшеничный 的专著).

第十六章　不动点原理

§1. Caccippoli-Banach原理

在这一章和下一章里，我们将主要考察非线性算子和非线性方程.

1.1. 对非线性方程的研究，我们从最简单的情形，即从推广的 Banach 逆算子定理开始.

考察完备度量空间 $\mathbf{X}$(不必是线性的)及其中的闭集合 Ω. 假定给定了一个从Ω到自身的算子 P. 如果存在点 $x^*\in\Omega$ 使得

$$x^*=P(x^*),$$

则称点 x^* 为算子 P 的不动点.

因此，算子 P 的不动点乃是方程

$$x=P(x) \tag{1}$$

的解.

算子 P 也可能没有不动点，例如，设 $\mathbf{X}=\Omega$ 是向量度量空间，则算子

$$P(x)=x+x_0 \qquad (x_0\neq\mathbf{0})$$

就不存在不动点. 但若 P 是压缩算子，即存在数 $\alpha<1$，使得对所有的 $x, x'\in\Omega$

$$\rho(P(x),P(x'))\leqslant\alpha\rho(x,x'), \tag{2}$$

则可保证不动点存在，而且还是唯一的. 亦即，下列定理成立.

定理 1. 如果 P 是压缩算子，则方程(1)在Ω中存在唯一的解 x^*.

这时，解 x^* 可作为序列$\{x_n\}$的极限而得到，其中

$$x_{n+1}=P(x_n) \qquad (n=0,1,2,\cdots),$$

而 x_0 是 Ω 中的任意元素.

序列 $\{x_n\}$ 收敛于解的速度由不等式

$$\rho(x_n, x^*)\leqslant\frac{\alpha^n}{1-\alpha}\rho(x_1, x_0) \qquad (n=0,1,2,\cdots) \tag{3}$$

给出.

证. 因为 $x_{n+1}=P(x_n)$, $x_n=P(x_{n-1})$,

故由(2)

$$\rho(x_{n+1}, x_n)\leqslant\alpha\rho(x_n, x_{n-1}).$$

逐次利用类似的不等式, 得到

$$\rho(x_{n+1}, x_n)\leqslant\alpha^n\rho(x_1, x_0).$$

所以

$$\rho(x_{n+p}, x_n)\leqslant\rho(x_{n+p}, x_{n+p-1})+\cdots+\rho(x_{n+1}, x_n)$$

$$\leqslant(\alpha^{n+p-1}+\cdots+\alpha^n)\rho(x_1, x_0)\leqslant\frac{\alpha^n}{1-\alpha}\rho(x_1, x_0). \tag{4}$$

由于 $\alpha^n\to0(n\to\infty)$, 故(4)式说明 $\{x_n\}$ 是自收敛的. 又因 $\mathbf{X}$ 是完备空间, 所以这个序列收敛于某 $x^*\in\mathbf{X}$. 但因 $x_n\in\Omega$, Ω 是闭集合, 所以 $x^*\in\Omega$, $P(x^*)$ 有意义. 由(2),

$$\rho(x_{n+1}, P(x^*))=\rho(P(x_n), P(x^*))\leqslant\alpha\rho(x_n, x^*)$$

$$(n=0,1,2,\cdots).$$

由于这个不等式的右端趋于零, 故

$$x_{n+1}\to P(x^*),$$

由此得到

$$x^*=P(x^*).$$

由(2)还可推证出解是唯一的. 事实上, 如果还存在一个解 $\tilde{x}\in\Omega$, 则就会有

$$\rho(\tilde{x}, x^*)=\rho(P(\tilde{x}), P(x^*))\leqslant\alpha\rho(\tilde{x}, x^*),$$

这只有当 $\rho(\tilde{x}, x^*)=0$ 才可能, 即 $\tilde{x}=x^*$.

最后，在(4)中令 $p\to\infty$ 而取极限，即得到(3)．

注．不等式(3)给出方程(1)可能成立的区域．特别，当 $n=0$ 时

$$\rho(x^*, x_0) \leqslant \frac{1}{1-\alpha}\rho(x_1, x_0). \tag{5}$$

1.2．条件(2)一般不能换成较弱的下列形式：

$$\rho(P(x), P(x')) < \rho(x, x') \qquad (x, x' \in \Omega, x \neq x'). \tag{6}$$

事实上，设 $\mathbf{X}=\mathbf{R}^1$ 是实数集合，$\Omega=\mathbf{X}$，而算子 P 用下式

$$P(x)=\frac{\pi}{2}+x-\operatorname{arc\,tg} x$$

来确定．不难看出算子 P 没有不动点，可是

$$\rho(P(x), P(x')) = |P(x)-P(x')| = |P'(\xi)|\,|x-x'|$$

$$= \left|\frac{\xi^2}{1+\xi^2}\right| |x-x'| < \rho\ (x, x'),$$

这里 ξ 是 x 和 x' 之间的一个点．但是，成立如下的定理．

定理 2．如果算子 P 将闭集 $\Delta\subset\Omega$ 变换为相对紧集 $\overline{\Delta}\subset\Omega$，而且对任何 x，$x'\in\Omega$ 条件(6)成立，则算子 P 存在唯一的不动点．

证．在紧集合 $\overline{\Delta}\subset\Omega$ 上考察算子 P．因为 P 显然连续，所以 $\varphi(x)=\rho(x, P(x))$ 是连续函数．该函数在紧集 $\overline{\Delta}$ 上的点 $x_0\in\overline{\Delta}$ 处取得最小值：

$$\rho(x_0, P(x_0)) = \min_{x\in\overline{\Delta}} \rho(x, P(x)).$$

设 $\rho(x_0, P(x_0))>0$．由(6)式可知

$$\rho(P(x_0), P^2(x_0)) < \rho(x_0, P(x_0)) = \min_{x\in\overline{\Delta}} \rho(x, P(x)). \tag{7}$$

但 $P(x_0)\in\overline{\Delta}$，所以与(7)式矛盾．因此 $\rho(x_0, P(x_0))=0$ 且 $x_0=P(x_0)$．

如果 $\bar{x}\in\Omega$ 是算子 P 的另一个不动点，则

$$\rho(\tilde{x}, x_0) = \rho(P(\tilde{x}), P(x_0)) < \rho(\tilde{x}, x_0),$$

从而得出矛盾. 定理证毕.

1.3. 在很多情形下，算子 P 都依赖于数值参数或其他参数. 此时，方程(1)的解同样依赖于这些参数. 可以证明，算子 P 对参数的连续依赖性导致了相应的方程的解对参数的连续依赖性. 我们来给出这一点的更确切的叙述.

设除了空间 $\mathbf{X}$ 外还有另一个度量空间 $\mathbf{Y}_0$. 假定对于每个 $y \in \mathbf{Y}_0$ 存在一个从 $\Omega \subset \mathbf{X}$ 到其自身的算子 P_y. 如果对任一序列 $\{y_n\} \subset \mathbf{Y}_0$, $y_n \to y_0$ 可知对每个 $x \in \Omega$

$$P_{y_n}(x) \to P_{y_0}(x), \tag{8}$$

则就说算子 P_y 在点 $y_0 \in \mathbf{Y}_0$ 处连续依赖于 y.

考察一族方程

$$x = P_y(x). \tag{9}$$

假定对于每个 $y \in \mathbf{Y}_0$, 方程(9)具有唯一的解. 显然，这个解将依赖于 y，所以可很自然地将这个解写成 x_y^*. 如果对任意的序列 $\{y_n\} \subset \mathbf{Y}_0$, 由 $y_n \to y_0$ 可推出 $x_{y_n}^* \to x_{y_0}^*$, 则称方程(9)的解当 $y = y_0$ 时连续依赖于 y.

如果对每个 $y \in \mathbf{Y}_0$, P_y 是压缩算子，则由算子 P_y 的连续性可推知方程(9)的解的连续性. 亦即，有如下定理.

定理 3. 如果对每个 $y \in \mathbf{Y}_0$, 算子 P_y 满足条件(2)，其中 α 不依赖于 y，并且在点 $y_0 \in \mathbf{Y}_0$ 算子 P_y 连续依赖于 y，则当 $y = y_0$ 时方程 (9) 的解连续依赖于 y.

证. 设 y 是 $\mathbf{Y}_0$ 中任意元素. 方程 (9) 的解 x_y^* 可作为序列 $\{x_n\}$:

$$x_{n+1} = P_y(x_n) \qquad (n = 0, 1, 2, \cdots; x_0 = x_{y_0}^*)$$

的极限而得. 由于 $x_{y_0}^* = P_{y_0}(x_{y_0}^*)$, 从(5)可得

$$\rho(x_y^*, x_{y_0}^*) \leqslant \frac{1}{1-\alpha}\rho(x_1, x_0) = \frac{1}{1-\alpha}\rho(P_y(x_{y_0}^*), P_{y_0}(x_{y_0}^*)).$$

注意到(8)，则容易看出 $y=y_0$ 时 x_y^* 的连续性.

注. 算子 P_y 可看成是将一对元素 (x, y) $(x\in\Omega, y\in\mathbf{Y}_0)$ 映射为元素 $P(x, y)=P_y(x)$ 的一个算子 P. 方程(9)的解 x_y^* 也可看成是将 $y\in\mathbf{Y}_0$ 映射为 $F(y)=x_y^*$ 的算子 F.

按照这样的看法，上面刚证明的定理 3 可改述为:

如果对每个 $y\in\mathbf{Y}_0$

$$\rho(P(x, y), P(x', y)) \leqslant \alpha\rho(x, x') \qquad (x, x'\in\Omega)$$

(其中 $\alpha<1$ 且不依赖于 y)，且对每个 $x\in\Omega$，算子 P 在点 $y_0\in\mathbf{Y}_0$ 处连续地依赖于 y，则算子 F 也在点 y_0 处连续.

上述形式的不动点原理是由 Banach[1] 和 Caccioppli 给出的. 关于这个原理在多值映射上的推广可参见 Иоффе 和 Тихомиров.

§2. 预 备 定 理

在下一节中我们将要建立第二不动点原理. 这个原理的证明是很难的，因为它需要借助于有限维空间中拓扑结构的细致结果. 我们在这一节中来给出这些取自拓扑学的预备知识.

2.1. 先引入一些在下面的讨论中要用到的新概念. 考察 Banach 空间 $\mathbf{X}$ 和其中的一组元素

$$x_0, x_1, x_2, \cdots, x_n. \tag{1}$$

假定差

$$x_1-x_0, x_2-x_0, \cdots, x_n-x_0 \tag{2}$$

线性独立. 我们构造元素(1)的凸包 $S(x_0, x_1, \cdots, x_n)$. 集合 $S(x_0, x_1, \cdots, x_n)$ 叫作张在顶点 $x_0, x_1, \cdots, x_n$ 上的单纯形，数 n 叫这个单纯形的维数.

注. 在上面关于单纯形的定义中，所有顶点并非处于平等地

位. 但是, 可以证明(请读者自行证明)实际上单纯形的所有顶点是地位平等的. 换言之,如果差(2)线性独立,则

$$x_0-x_k, \cdots, x_{k-1}-x_k, x_{k+1}-x_k, \cdots, x_n-x_k \quad (k=1, 2, \cdots, n)$$

中每一个也必是线性独立的.

设 $S(x_0, x_1, \cdots, x_n)$是一单纯形, 从中取出某 k 个不同的顶点 $x_{i_1}, x_{i_2}, \cdots, x_{i_k}$, 并以这 k 个点为顶点构成一个$(n-k)$维的单纯形(由上面的注, 这是可以作到的). 所构成的这个新的单纯形叫做单纯形 $S(x_0, x_1, \cdots, x_n)$的以 $x_{i_1}, x_{i_2}, \cdots$, x_{i_k} 为顶点的 $(n-k)$ 维边界.

再设 $S=S(x_0, x_1, \cdots, x_n)$是某个单纯形, 每个元素$x\in S$ 可表示成

$$x=\alpha_0x_0+\alpha_1x_1+\cdots+\alpha_nx_n, \tag{3}$$

其中系数 α_k 满足如下关系式:

$$\alpha_0+\alpha_1+\cdots+\alpha_n=1, \alpha_k\geqslant 0 \qquad (k=0, 1, 2, \cdots, n).$$

今证表达式(3)是唯一的. 事实上,由于 $\alpha_0=1-\sum_{k=1}^{n}\alpha_k$, 由(3)可得

$$x=x_0+\sum_{k=1}^{n}\alpha_k(x_k-x_0). \tag{4}$$

假定除(3)外还有另一个这样的表达式

$$x=\alpha_0'x_0+\alpha_1'x_1+\cdots+\alpha_n'x_n,$$

类似地可得

$$x=x_0+\sum_{k=1}^{n}\alpha_k'(x_k-x_0).$$

将这个等式代入(4),则

$$\sum_{k=1}^{n}\alpha_k(x_k-x_0)=\sum_{k=1}^{n}\alpha_k'(x_k-x_0),$$

由于差(2)的线性独立性,上式仅当

$$\alpha_1=\alpha_1', \alpha_2=\alpha_2', \cdots, \alpha_n=\alpha_n', \quad \alpha_0'=\alpha_0$$

时才有可能. 这就证明了(3)的唯一性.

我们将单值地确定了元素 x 的这组数 $\alpha_0, \alpha_1, \cdots, \alpha_n$ 叫做点 $x\in S$ 的对称坐标*).

应指出的是,对于与顶点 $x_{i_1}, x_{i_2}, \cdots, x_{i_k}$ 相对的边界点来说,有 $\alpha_{i_1}=\alpha_{i_2}=\cdots=\alpha_{i_k}=0$.

所有坐标都是正数的点叫单纯形的内点. 所有坐标彼此均相等的点叫单纯形的中心.

我们来引进关于单纯形 $S(x_0, x_1, x_2, \cdots, x_n)$ 的子分划的概念. 所谓单纯形的子分划就是有限个与所给单纯形维数相同的单纯形之和,它由下列归纳形式来确定:

对一维单纯形 $S(x_0, x_1)$ 来说,我们称由单纯形 $S\left(x_0, \frac{1}{2}(x_0+x_1)\right)$和单纯形 $S\left(\frac{1}{2}(x_0+x_1), x_1\right)$的和集构成其子分划. 假定我们已对维数小于 n 的所有单纯形都定义了子分划, 我们再来考察 n 维单纯形 $S=S(x_0, x_1, \cdots, x_n)$. 取单纯形 S 的某个 $(n-1)$ 维边界. 由于这个边界是$(n-1)$维单纯形, 所以对它是已经定义了子分划的. 设这个子分划是由 m 个$(n-1)$维单纯形 $S_1, S_2, \cdots, S_m$构成. 将 n 维单纯形 S 的中心 x^* 补充到 S_k $(k=1, 2, \cdots, m)$ 的顶点中去, 一共就有$(n+1)$个点. 我们在这$(n+1)$个点上构造一个 n 维单纯形(这样做的可能性请读者自行验证). 对已知单纯形的每个$(n-1)$维边界都进行类似的构造,即得其子分划.

不难看出,所构造的子分划没有公共内点,这些子分划的交是

*) 数 $\alpha_0, \alpha_1, \cdots, \alpha_n$ 也叫点 $x\in S$ 的重心坐标. 因若在每个顶点 x_k 上置放 质量 α_k,则所得到的质点系的惯性中心就在点 x 处.

公共边界.

我们来考察某个单纯形 S 和它的子分划. 对于 S 的每个部分单纯形再作子分划, 就得到 S 的二重子分划, 继续上述过程就得到多重子分划.

易见, 随着子分划的重数的增加, 部分单纯形的半径趋于零.

2.2. 现在引进一些引理, 我们在后面将要根据这些引理来证明第二不动点原理.

设 $S=S(x_0, x_1, \cdots, x_n)$ 是一个 n 维单纯形, 它的子分划(任意重的)由单纯形

$$S_1, S_2, \cdots, S_m \tag{5}$$

组成.

令(5)中的每个单纯形的每个顶点 z 对应于数 $0, 1, \cdots, n$ 中的一个数, 即在(5)中单纯形所有顶点的集合上定义一个取值 $0, 1, 2, \cdots, n$ 的函数 ν. 这样, (5)中每个单纯形对应于一组数 $\nu(S_j)$, 这是由对应于此单纯形顶点的 $(n+1)$ 个数中的一些数组成的. 如果 $\nu(S_j)=\{0, 1, 2, \cdots, n\}$, 则称这样的单纯形 S_j 是正规的. 这里认为 $\nu(S_j)$ 中数的排列次序是无关紧要的.

一般来说, 当 ν 是任意函数时, 正规单纯形可能不存在. 但是, 有下面的引理.

引理 1. 设函数 ν 满足这样的条件: 若 z 是(5)中某个单纯形的顶点, 这个顶点属于基本单纯形 S 的 k 维边界 $S(x_{i_0}, x_{i_1}, \cdots, x_{i_k})$, 则 $\nu(z)$ 只能取 $i_0, i_1, \cdots, i_k$ 中的一个数.

这时, 单纯形(5)中至少有一个是正规的.

证. 我们先证明正规单纯形的个数是奇数. 如果单纯形 S 的维数等于零, 也就是说如果单纯形退化成一个点, 则结论是显然的. 设对 $(n-1)$ 维单纯形已经证明结论成立, 今证对 n 维单纯形结论也成立. 考察单纯形 S_k 的 $(n-1)$ 维边界 $S_k^{(i)}$. 如果 $\nu(S_k^{(i)})=$

$\{0, 1, 2, \cdots, n\}$，则称 $S_k^{(i)}$ 是划出的．如果单纯形 S_k 是正规的，则容易验证它的划出边界的数 $\rho_k=1$；若单纯形 S_k 不是正规的，则 $\rho_k=2$．因此，和式 $\rho=\sum_{k=1}^{m}\rho_k$ 和所有正规单纯形的数目具有相同的奇偶性．我们将单纯形(5)的划出边界分成两组．第一组中不包括基本单纯形 S 的任何一个 $(n-1)$ 维边界．这样的划出边界的数目用 ρ' 表示．由于每个这种划出边界正好是(5)中两个单纯形的公共边界，所以在计算和 ρ 时，这种边界要计算两次．第二组边界是完全包含在单纯形 S 的一个 $(n-1)$ 维边界之中的划出边界．因为这种边界只能是(5)中一个单纯形的边界，所以这种边界的数目 $\rho''=\rho-2\rho'$．利用引理的条件不难确信：所有第二组划出边界都包含在单纯形 S 的边界 $S'=S(x_0, x_1, \cdots, x_{n-1})$ 之中．但所有可能落在 S' 中的单纯形(5)的 $(n-1)$ 维边界构成单纯形 S' 的子分划，并且函数 ν（仅看作是对所得到的子分划的函数）满足引理的条件．因而，由归纳假定，所指出的子分划中正规单纯形的个数是奇数．注意这时正规单纯形正好是第二组划出边界，所以 ρ'' 是奇数，因而 $\rho=2\rho'+\rho''$ 亦是奇数．

引理于是得证．

引理 2. 设 $S=S(x_0, x_1, \cdots, x_n)$ 是 n 维Banach空间中的 n 维单纯形而且存在这样的闭集合 $F_0, F_1, \cdots, F_n$ 使得对不论怎样的下标组 $i_0, i_1, \cdots, i_k\ (k=0, 1, \cdots, n)$ 均有

$$S(x_{i_0}, x_{i_1}, \cdots, x_{i_k})\subset\bigcup_{m=0}^{k}F_{i_m}.$$

这时，交 $\bigcap_{k=0}^{n}F_k$ 是非空集．

证．考察单纯形 S 的任意重子分划．设 z 是构成这种子分划的单纯形中某个单纯形的顶点而 $S(x_{i_0}, \cdots, x_{i_k})$ 是包含 z 的边界．

根据引理的条件，$z\in\bigcup\limits_{m=0}^{k}F_{i_m}$. 假定 $z\in F_{i_s}$，这时令 $\nu(z)=i_s$. 显然，这样的函数 ν 满足引理 1 的条件，因而存在这样的正规部分单纯形，这个单纯形与集合 F_j 的每一个的交都是非空的（包含顶点中的一个）.

现在来考察 p 重子分划序列 $(p=1,2,\cdots)$. 对于第 p 个子分划，存在点 $z_0^{(p)}, z_1^{(p)}, \cdots, z_n^{(p)}$（正规部分单纯形的顶点）使得

$$z_k^{(p)}\in F_k \qquad (k=0,1,\cdots,n;\ p=1,2,\cdots). \tag{6}$$

由于集合 S 的紧性，可指出这样的上升自然数序列 $\{p_j\}$ 使得序列 $\{z_0^{(p_j)}\}$ 收敛（设其极限为 z^*）. 由于单纯形 S 的闭性，则必 $z^*\in S$. 因当 $p\to\infty$ 时部分单纯形的直径趋于零，所以

$$z_k^{(p_j)}\longrightarrow z^* \qquad (k=0,1,\cdots,n).$$

由此，考虑到集合 F_k 的闭性以及(6)，则

$$z^*\in F_k \qquad (k=0,1,2,\cdots,n),$$

这就证明了引理.

2.3. 下面的引理就已经给出了某种形式的不动点原理（Brouwer 原理），不过是在相当受约束的条件下给出来的.

引理 3. 将 Banach 空间 $\mathbf{X}$ 中的 n 维单纯形 S 映射到自身的连续算子 P 具有不动点.

证. 设 $\alpha_0,\alpha_1,\cdots,\alpha_n$ 是任意点 $x\in S$ 的对称坐标. 令

$$\beta_j=f_j(\alpha_0,\alpha_1,\cdots,\alpha_n) \qquad (j=0,1,\cdots,n)$$

为点 $P(x)$ 的对称坐标. 由于 $P(x)\in S$，所以这样的表示是有意义的. 显然

$$\beta_0+\beta_1+\cdots+\beta_n=1,\ \beta_j\geqslant 0 \qquad (j=0,1,\cdots,n).$$

由算子 P 的连续性不难推知函数 f_j 都是连续的. 用 $F_j(j=0,1,\cdots,n)$ 表示其对称坐标满足关系

$$\alpha_j\geqslant\beta_j$$

的 $x\in S$ 的集合. 由于函数 f_j 的连续性,集合 F_j 是闭的. 我们来证明集合 F_j 满足引理 2 的条件.

事实上,设 $x\in S(x_{i_0}, x_{i_1}, \cdots, x_{i_k})$. 假设 $x\notin \bigcup_{m=0}^{k} F_{i_m}$. 这表示点 x 和 $P(x)$ 的对称坐标 $\alpha_0, \alpha_1, \cdots, \alpha_n$ 和 $\beta_0, \beta_1, \cdots, \beta_n$ 有关系

$$\alpha_{i_m} < \beta_{i_m} \qquad (m=0, 1, \cdots, k).$$

因为对于 $i\neq i_m (m=0, 1, \cdots, k)$ 应有 $\alpha_i=0$, 因而对这样的下标 $\alpha_i \leqslant \beta_i$, 我们就得到了

$$\sum_{i=0}^{n} \alpha_i < \sum_{i=0}^{n} \beta_i,$$

而这是不可能的,因为这两个和式都应等于 1.

因此,由引理 2, 可找到 $x^*\in \bigcap_{k=0}^{n} F_k$. 如果 $\alpha_0^*, \alpha_1^*, \cdots, \alpha_n^*$ 和 $\beta_0^*, \beta_1^*, \cdots, \beta_n^*$ 分别是点 x^* 和 $P(x^*)$ 的对称坐标,则

$$\alpha_j^* \geqslant \beta_j^* \qquad (j=0, 1, \cdots, n).$$

由此,

$$\sum_{i=0}^{n} \alpha_i^* \geqslant \sum_{i=0}^{n} \beta_i^*. \tag{7}$$

(7) 式中的等号仅当

$$\alpha_i^* = \beta_i^* \qquad (i=0, 1, \cdots, n) \tag{8}$$

时成立. 但 (7) 中的两个和式均等于 1, 因而 (8) 式成立, 这就是说 $x^* = P(x^*)$.

引理得证.

2.4. 我们现在将引理 3 的结果转移到有限维空间的任意闭有界凸集合上去. 为此,还要先证明一个引理.

记得在 III. 2.3 中我们曾称 Banach 空间 **X** 中具有非空内部

的闭凸集Ω为凸体.

引理4. 设Ω是 Banach 空间$\mathbf{X}$中的按范数有界凸体而$B_{\mathbf{X}}$是空间$\mathbf{X}$中的闭单位球. 这时,集合Ω与$B_{\mathbf{X}}$是同胚的.

证. 不失一般性, 可以认为$\mathbf{0}$是Ω的内点而$B_{\mathbf{X}}\subset\Omega$. 这时$\Omega$是$\mathbf{0}$邻域,我们可以考察它的 Minkowski 泛函$p$(II. 3. 2).

由引理 III. 2. 1 证明的 Minkowski 泛函的一般性质得知p是连续的并且$\Omega=\{x\in\mathbf{X}:p(x)\leqslant 1\}$. 如果$x$是$\mathbf{X}$中任意非零元素,则$\frac{x}{\|x\|}\in B_{\mathbf{X}}\subset\Omega$. 因而$p\left(\frac{x}{\|x\|}\right)\leqslant 1$,即有不等式

$$p(x)\leqslant\|x\|,\tag{9}$$

这个不等式当x为零元时自然也成立.

其次,由于集合Ω的有界性,存在一个以零为中心以$r>0$为半径而包含集合Ω的开球K_r. 因为对于任意的$x\neq\mathbf{0}$, $\frac{rx}{\|x\|}\notin K_r$, 故$\frac{rx}{\|x\|}\notin\Omega$, 因而$p\left(\frac{rx}{\|x\|}\right)>1$. 由此得到

$$p(x)\geqslant\frac{1}{r}\|x\|\qquad(x\in\mathbf{X}).\tag{10}$$

我们来定义映射T:

$$T(x)=x\frac{p(x)}{\|x\|}\qquad(x\neq\mathbf{0}),\qquad(T(\mathbf{0})=\mathbf{0}),\qquad(x\in\Omega).\tag{11}$$

因为集合Ω与所有使得$p(x)\leqslant 1$的$x\in\mathbf{X}$所组成的集合重合,所以有$T(\Omega)=B_{\mathbf{X}}$. 此外显然可看出映射T是一对一的, 逆算子T^{-1}由下式给出:

$$T^{-1}(x)=x\frac{\|x\|}{p(x)}\quad(x\neq\mathbf{0});\qquad T^{-1}(\mathbf{0})=\mathbf{0}.\tag{12}$$

映射T在点$x_0\neq\mathbf{0}$处的连续性可由泛函p的连续性推出来.

如果 $x_n \to \mathbf{0}$，因 $\left\|\frac{x_n}{\|x_n\|}\right\|=1$ 而 $p(x_n)\to\mathbf{0}$，所以 $T(x_n)\to\mathbf{0}=T(\mathbf{0})$.

可完全类似地来验证 T^{-1} 的连续性．如果 $x_n\to x_0\neq\mathbf{0}$，则由于 $p(x_0)\geqslant\frac{1}{r}\|x_0\|>0$（见(10)式)，由泛函 p 的连续性可推出 $T^{-1}(x_n)\to T^{-1}(x_0)$．当 $x_0=\mathbf{0}$ 时，再由(10)，

$$\|T^{-1}(x_n)\|=\frac{\|x_n\|}{p(x_n)}\|x_n\|\leqslant\frac{r\|x_n\|}{\|x_n\|}\|x_n\|=r\|x_n\|\to 0,$$

亦即 $T^{-1}(x_n)\to\mathbf{0}=T^{-1}(\mathbf{0})$.

引理证毕.

推论．n 维 Banach 空间中的有界凸体与 n 维单纯形同胚.

事实上，n 维单纯形是 n 维空间中的凸体，它和 Ω 同时与球 B 同胚.

现在，有了引理 3 以及引理 4 的推论，我们可以来证明基本引理了.

引理 5．设 Ω 是 n 维 Banach 空间 $\mathbf{X}$ 的有界凸体．将集合 Ω 映射到自身的连续算子 P 存在不动点.

证．在空间 $\mathbf{X}$ 中取线性独立元素 $x_1, x_2, \cdots, x_n$．设 $x_0=\mathbf{0}$，取张在顶点 $x_0, x_1, \cdots, x_n$ 上的单纯形 S．由引理 4 的推论，存在将 S 映到体 Ω 上的一对一的和双连续的映射 T．考察算子

$$P_0=T^{-1}PT.$$

算子 P_0 连续并将单纯形 S 映为自身．因而，由引理3，存在不动点 x_0:

$$x_0=P(x_0).$$

因为

$$P(x^*)=PT(x_0)=TT^{-1}PT(x_0)=TP_0(x_0)=T(x_0)=x^*,$$

所以点 $x^*=T(x_0)$ 是算子 P 的不动点。

关于将单纯形变为自身的连续变换的不动点定理是由 Brouwer [1]在联系到维数不变性时建立的. 但上述的证明方法则是属于 Kuratowski.

§3. Schauder原理

3.1. 由上节的结果我们可以建立第二不动点原理，即Schauder 原理.

定理1. 将 Banach 空间X的凸紧集Ω映射到自身的连续算子P具有不动点.

证. 任取$\varepsilon>0$. 因Ω是紧集，故在Ω中可求得有限的ε-网，设这个ε-网由元素

$$x_1, x_2, \cdots, x_m \tag{1}$$

组成.

构造元素(1)的凸包Ω_0. 显然$\Omega_0\subset\Omega$. 这时集合Ω_0具有有限维数$n\leqslant m-1$*).

借助于关于维数的归纳法不难证明，集合Ω_0可表示成这样的n维单纯形的和，使得

a) (1) 的所有点都是这些单纯形的顶点;

b) 两个单纯形要么没有公共点，要么它们的交是k维公共边界($k=0,1,\cdots,n-1$).

我们对所构成的每个单纯形作如此高重的子分划，使得由这种子分划而得到的所有部分单纯形的直径均小于ε. 用

$$S_1, S_2, \cdots, S_p \tag{2}$$

表示所构成的这些单纯形. 显然，单纯形(2)所有顶点的集合也构

*) 我们说集合$E\subset X$的维数等于n，乃是指存在元素$\bar{x}_0, \bar{x}_1, \cdots, \bar{x}_n$使得 1) 差$\bar{x}_k-\bar{x}_0(k=1,2,\cdots,n)$线性独立, 2) 对每个$x\in E$，可表示成$x=\bar{x}_0+\sum_{k=1}^{n}a_k(\bar{x}_k-\bar{x}_0)$，并且对任意的$k=1,2,\cdots,n$存在$x$使得对应的系数$a_k\neq0$.

成 ε-网. 还应指出的是单纯形(2)满足条件 a)和 b).

现在来看算子 P. 由集合 Ω 的紧性, 从 P 的连续性推知它是一致连续的, 亦即对任何 $\varepsilon>0$ 可以指出这样的 $\delta>0$ 使得由 $\|x-x'\|<\delta(x, x'\in\Omega)$ 可得到

$$\|P(x)-P(x')\|<\varepsilon \qquad (x, x'\in\Omega). \tag{3}$$

如果需要, 可对单纯形(2)作子分划, 我们就能使分划后诸单纯形的直径全小于 δ. 可以假定这样的子分划已经作了, 即所有的单纯形(2)的直径不仅小于 ε 而且都小于 δ.

现在我们来构造将 Ω_0 映射为自身的算子 P_ε, 它是算子 P 的单纯形近似. 先在单纯形(2)的顶点处定义算子 P_ε. 设 z 是单纯形 (2) 中某个单纯形的某一顶点. 由于 $P(z)\in\Omega$ 而单纯形(2)的顶点全体构成 ε-网, 所以可以求得顶点 $\bar{z}$ 使得 $\|\bar{z}-P(z)\|<\varepsilon$. 我们就命 $\bar{z}=P_\varepsilon(z)$.

现在设 $x\in\Omega_0$ 不是单纯形(2)中任何一个的顶点. 假定 $x\in S_k$. 用 $x_0^{(k)}, x_1^{(k)}, \cdots, x_n^{(k)}$ 表示单纯形 S_k 的顶点, 并且将 x 表示成

$$x=\sum_{i=0}^{n}\alpha_i^{(k)}x_i^{(k)} \qquad \left(\sum_{i=0}^{n}\alpha_i^{(k)}=1; \alpha_i^{(k)}\geqslant 0, i=0, 1, \cdots, n\right)$$

的形式. 这时, 我们就命

$$P_\varepsilon(x)=\sum_{i=0}^{n}\alpha_i^{(k)}P_\varepsilon(x_i^{(k)}). \tag{4}$$

如果包含点 x 的单纯形是唯一的, 这个定义并不需要补充说明. 可是, 如果 $x\in S_r(r\neq k)$, 则还应该来证明 $P_\varepsilon(x)$ 不依赖于单纯形的选择方式. 根据条件 b), 单纯形 S_k 和 S_r 的交是其公共边界. 我们假定这个公共边界是张在单纯形 S_k 的顶点 $x_{i_0}^{(k)}, x_{i_1}^{(k)}, \cdots, x_{i_s}^{(k)}$ 上; 用 $x_0^{(r)}, x_1^{(r)}, \cdots, x_n^{(r)}$ 表示单纯形 S_r 的顶点, 可以认为

$$x_{i_j}^{(k)}=x_{i_j}^{(r)} \qquad (j=1, 2, \cdots, s). \tag{5}$$

将 x 表示为

$$x=\sum_{i=1}^{n}\alpha_i^{(r)}x_i^{(r)} \qquad \left(\sum_{i=1}^{n}\alpha_i^{(r)}=1;\alpha_i^{(r)}\geqslant 0,\ i=0,1,\cdots,n\right).$$

注意到 2.1 中的注，则

$$\alpha_i^{(k)}=\alpha_i^{(r)}=0 \qquad (i\neq i_j;\ j=1,2,\cdots,s). \tag{6}$$

由(5)

$$\sum_{j=0}^{s}\alpha_{i_j}^{(k)}x_{i_j}^{(k)}=\sum_{j=0}^{s}\alpha_{i_j}^{(r)}x_{i_j}^{(k)},$$

由此推出等式

$$\alpha_{i_j}^{(k)}=\alpha_{i_j}^{(r)} \qquad (j=0,1,2,\cdots,s). \tag{7}$$

因此，如果我们从单纯形 S_r 出发来定义 $P_\varepsilon(x)$，则由(5)，(6)及(7)就有

$$P_\varepsilon(x)=\sum_{i=0}^{n}\alpha_i^{(r)}P_\varepsilon(x_i^{(r)})=\sum_{j=0}^{s}\alpha_{i_j}^{(r)}P_\varepsilon(x_{i_j}^{(r)})$$
$$=\sum_{j=0}^{s}\alpha_{i_j}^{(k)}P_\varepsilon(x_{i_j}^{(k)})=\sum_{i=0}^{n}\alpha_i^{(k)}P_\varepsilon(x_i^{(k)}),$$

这正是所要证明的.

通过不太复杂的讨论可以证明所建立的算子 P_ε 在 Ω_0 上连续而且将 Ω_0 变为自身.

因此，对算子 P_ε 来说，上一节引理 5 的条件成立. 因而算子 P_ε 存在不动点 $x_\varepsilon\in\Omega_0$:

$$x_\varepsilon=P_\varepsilon(x_\varepsilon).$$

设 $z_0, z_1, \cdots, z_n$ 是单纯形(2)中点 x_ε 所属的那个单纯形的顶点. 因 $\|z_i-z_j\|<\delta$，所以按(3)有

$$\|P(z_i)-P(z_j)\|<\varepsilon \qquad (i,j=0,1,\cdots,n).$$

另外，按照定义

$$\|P_\varepsilon(z_i)-P(z_j)\|<\varepsilon \qquad (i=0,1,\cdots,n). \tag{8}$$

所以，我们得到

$$\|P_\varepsilon(z_i)-P_\varepsilon(z_j)\|<3\varepsilon \quad (i, j=0, 1, 2, \cdots, n). \tag{9}$$

其次，将 x_ε 表示成

$$x_\varepsilon=\sum_{i=0}^{n}\alpha_i^{(\varepsilon)}z_i \qquad \left(\sum_{i=0}^{n}\alpha_i^{(\varepsilon)}=1;\quad \alpha_i^{(\varepsilon)}\geqslant 0;\quad i=0, 1, \cdots, n\right),$$

按照算子 P_ε 的定义

$$x_\varepsilon=P_\varepsilon(x_\varepsilon)=\sum_{i=0}^{n}\alpha_i^{(\varepsilon)}P_\varepsilon(z_i).$$

由此，再考虑到(9)，我们就得到

$$\|x_\varepsilon-P_\varepsilon(z_j)\|=\left\|\sum_{i=0}^{n}\alpha_i^{(\varepsilon)}[P_\varepsilon(z_i)-P_\varepsilon(z_j)]\right\|$$

$$\leqslant 3\varepsilon\sum_{i=0}^{n}\alpha_i^{(\varepsilon)}=3\varepsilon \qquad (j=0, 1, 2, \cdots, n). \tag{10}$$

最后，由关系式

$$\|x_\varepsilon-z_i\|<\delta \qquad (i=0, 1, \cdots, n)$$

以及(3)推知

$$\|P(x_\varepsilon)-P(z_i)\|<\varepsilon \qquad (i=0, 1, \cdots, n). \tag{11}$$

利用(10)，(8)及(11)，最后得到

$$\|x_\varepsilon-P(x_\varepsilon)\|\leqslant\|x_\varepsilon-P_\varepsilon(z_i)\|+\|P_\varepsilon(z_i)-P(z_i)\|$$
$$+\|P(z_i)-P(x_\varepsilon)\|<5\varepsilon.$$

这也就是说，对任意的 $\varepsilon>0$ 都存在点 $x_\varepsilon\in\Omega$，使得

$$\|x_\varepsilon-P(x_\varepsilon)\|<5\varepsilon.$$

取序列 $\varepsilon_k\to 0$，对每个 $k=1, 2, \cdots$ 得点 $x_k=x_{\varepsilon_k}$。由于集合 Ω 的紧性和闭性，不失一般性，可以认为

$$x_k\xrightarrow[k\to\infty]{}x^*\in\Omega.$$

但这时

$$\|x^*-P(x^*)\| \leqslant \|x^*-x_k\| + \|x_k-P(x_k)\| + \|P(x_k)-P(x^*)\|,$$

而不等式右端三项均趋于零（最后一项趋于零是由于算子P的连续性）. 因而

$$x^*=P(x^*),$$

这就证明了不动点的存在性.

注. 我们提请读者注意，定理的两个条件（集合Ω的凸性和紧性）是本质性的. 确实，不难构造出这样的例子，它表明紧性条件不能减弱，不能用集合Ω的有界性来代替紧性.

3.2. 利用 Banach 空间中相对紧集的闭凸包还是紧集（定理 III.2.3 的推论）这一结论可将定理 1 叙述成更一般的形式.

定理2. 映 Banach 空间 X 中的闭凸集Ω到紧集$\Delta\subset\Omega$的连续算子P存在不动点.

证. 考察集合Δ的闭凸包$\Omega_0=\overline{\mathrm{co}}(\Delta)$. 显然$\Omega_0\subset\Omega$. 而且$\Omega_0$是凸的紧集.

算子P将Ω_0映到Δ内，由于$\Delta\subset\Omega_0$，所以P将Ω_0映射到自身. 因此，若将P仅看作是Ω_0上的算子，则可看到P完全满足定理 1 的条件. 从而应用定理 1 即完成了证明.

对定理 1 和定理 2 的证明作相应的改变可以证明当 X 是局部凸空间时类似的不动点存在定理（见 Dunford 和 Schwartz-I）.

定理 1 和定理 2 形式的不动点原理是在 Schauder[1] 的文章中建立起来的. 更深刻的不动点存在定理，特别是对局部凸空间情形是由 Leray 证明的（参见 Leray 和 Schauder[1]）. 后来，这些问题在 M. A. Красносельский 的工作（见 Красносельский-I, Красносельский[1]—[3]）中得到了明显的推进.

§ 4. 不动点原理的应用

我们来考察以上所证明的定理的某些应用.

4.1. 设 $\Phi(s,u)$ 是定义在带形区域 $a\leqslant s\leqslant b$, $-\infty<u<\infty$ 上的两个实变数的函数, 假定 Φ 关于 u 连续并存在关于 u 的满足条件

$$0<m\leqslant \Phi'_u(s,u)\leqslant M \qquad (a\leqslant s\leqslant b;\ -\infty<u<\infty) \tag{1}$$

的连续导数 Φ'_u. 在这些条件之下, 存在唯一的在 $[a,b]$ 上连续的函数 $u=x^*(s)$, 满足方程

$$\Phi(s,x^*(s))=0. \tag{2}$$

这个结果可以借助于定理 1.1 而得到. 为此, 我们在空间 $\mathbf{X}=\mathbf{C}[a,b]$ 中考察由下述形式所定义的算子 P:

$$y=P(x),\ y(s)=x(s)-\frac{2}{M+m}\Phi(s,x(s)) \qquad (s\in[a,b]).$$

不难验证 P 是压缩算子. 事实上, 如果 $y=P(x)$, $\tilde{y}=P(\tilde{x})$, 则对任何 $s\in[a,b]$ 有

$$\begin{aligned}|y(s)-\tilde{y}(s)|&=\Big|x(s)-\tilde{x}(s)-\frac{2}{M+m}[\Phi(s,x(s))\\&\qquad-\Phi(s,\tilde{x}(s))]\Big|\\&=|x(s)-\tilde{x}(s)|\left|1-\frac{2}{M+m}\Phi'_u(s,\theta(s))\right|\\&\leqslant\alpha\|x-\tilde{x}\|,\end{aligned}$$

其中

$$\alpha=\frac{M-m}{M+m},$$

由 (1) 可知 $\alpha<1$.

所以算子 P 在 $\mathbf{C}[a,b]$ 中存在唯一的不动点 x^*. 剩下就只要

说明关系式 $x^*=P(x^*)$ 与等式(2)是等价的了.

4.2. 其次，考察常微分方程组

$$\frac{dx_i}{dt}=\varphi_i(x_1, x_2, \cdots, x_n; t) \qquad (i=1, 2, \cdots, n). \tag{3}$$

如果引进 n 维向量函数，它由 n 个实函数的有序组构成，则可将方程组(3)写作一个方程

$$\frac{dx}{dt}=\varphi(x, t) \tag{4}$$

的形式，其中 $x(t)$ 是以 $x_1(t), x_2(t), \cdots, x_n(t)$ 为分量的向量函数，$\varphi(x, t)$ 则是以 $\varphi_1(x_1, \cdots, x_n; t), \cdots, \varphi_n(x_1, \cdots, x_n; t)$ 为分量的向量函数，导数 $\frac{dx}{dt}$ 看成是以 $\frac{dx_1}{dt}, \frac{dx_2}{dt}, \cdots, \frac{dx_n}{dt}$ 为分量的函数.

假定我们要求出方程组(1)在区间 $[a, b]$ 上满足初始条件

$$x_i(a)=y_0^{(i)} \qquad (i=1, 2, \cdots, n) \tag{*}$$

的解. (*)式写成向量形式就是

$$x(a)=y_0, \tag{5}$$

其中 y_0 是空间 $\mathbf{R}^n$ 中的元素，它的坐标是 $y_0^{(1)}, y_0^{(2)}, \cdots, y_0^{(n)}$. 这个问题的解等价于积分方程

$$x(t)=y_0+\int_a^t \varphi(x(\tau), \tau)d\tau \tag{6}$$

的解.

定理1. 设向量函数 $\varphi(y, t)$ 对于

$$y\in\mathbf{R}^n, \ |y-y_0|<\delta^{*)}, \ t\in[a, b']$$

是有定义的. 此外，还设对如上的 y, t 函数 $\varphi(y, t)$ 有界，对每个可能的 y 值 φ 关于 t 可测并满足关于 y 的 Lipschitz条件:

$$|\varphi(y, t)-\varphi(\tilde{y}, t)|\leqslant K|y-\tilde{y}|, \tag{7}$$

---

*) 这里和以后均以 $|y|$ 表示向量 y 的欧氏长度.

其中 K 不依赖于 t.

记 $M=\sup|\varphi(y,t)|\qquad(|y-y_0|\leqslant\delta,\ t\in[a,b'])$,
这时，如果区间 $[a,b]\subset[a,b']$ 充分小，也就是说如果

$$b-a<\min\left[\frac{\delta}{M},\frac{1}{K}\right],\tag{8}$$

则积分方程(6)(因而也就是微分方程组(3))有唯一的满足给定初始条件的解. 而且解连续依赖于初始向量 y_0.

证. 我们引进由 $[a,\ b]$ 上所有连续的 n 维向量函数组成的空间 $\mathbf{C}_n([a,b],\mathbf{R}^n)$. 如果依自然的方式将 $\mathbf{C}_n$ 线性化，并设

$$\|x\|=\max_{t\in[a,b]}|x(t)|,\quad(x\in\mathbf{C}_n),$$

则 $\mathbf{C}_n$ 成为 Banach 空间.

考虑到要对方程(6)应用定理 1.3，所以我们设 $\mathbf{X}=\mathbf{C}_n$，$\mathbf{Y}=\mathbf{R}^n$. 方程(6)可写成形如

$$x=P_{y_0}(x)\tag{9}$$

的泛函方程，其中 P_y 是定义在集合 Ω 上的积分算子:

$$z=P_y(x)\qquad(y\in\mathbf{R}^n,\ x\in\mathbf{C}_n),$$

$$z(t)=y+\int_a^t\varphi(x(\tau),\tau)d\tau,\tag{10}$$

而 Ω 则是由满足条件

$$|x(t)-y_0|\leqslant\delta\qquad(t\in[a,b])$$

的所有 $x\in\mathbf{C}_n$ 组成的集合. 作为 $\mathbf{Y}_0$ 应取满足条件

$$|y-y_0|<\delta_1=\delta-M(b-a)$$

的向量 $y\in\mathbf{R}^n$ 的全体.

算子 P_y 对于每个 $y\in\mathbf{Y}_0$ 将 Ω 变到自身. 事实上，

$$\begin{aligned}|z(t)-y_0|&\leqslant|y-y_0|+\left|\int_a^t\varphi(x(\tau),\ \tau)d\tau\right|\\&\leqslant\delta_1+M(b-a)=\delta.\end{aligned}$$

还要验证算子 P_y 满足 §1 中的条件(2). 设 $x,\tilde{x}\in\Omega$, $y\in\mathbf{Y}_0$,

$$\|P_y(x)-P_y(\tilde{x})\|=\max_{t\in[a,b]}\left|\int_a^t[\varphi(x(\tau),\tau)-\varphi(\tilde{x}(\tau),\tau)]d\tau\right|$$
$$\leqslant K(b-a)\|x-\tilde{x}\|.$$

记

$$\alpha=K(b-a),$$

按(8)式，有 $\alpha<1$.

算子 P_y 对 y 的连续性是显然的.

于是，按照定理 1.3，方程(9)(因而也就是积分方程(6))有唯一的解，而且该解连续依赖于 y.

注. 方程组(3)的解 x_y^* 当 $y=y_0$ 时对初始向量 y 的连续依赖性在这种情形下意味着：对每一 $\varepsilon>0$ 都可找到这样的 $\eta>0$ 使得只要 $|y-y_0|<\eta$ 就有 $\|x_y^*-x_{y_0}^*\|_C<\varepsilon$.

4.3. 对算子(9)应用 Schauder 原理可以在对函数 $\varphi(y,t)$ 作更弱的假定之下建立方程(6)(因而也就是方程组(1))的解的存在性. 但这时不能证明解的唯一性，特别是不能证明解对初始向量的连续依赖性.

定理 2. 设函数 φ 除了条件(7)之外满足定理 1 的所有条件，而将条件(7)用一个较弱的要求来代替：对每个 $t\in[a,b']$，函数 $\varphi(y,t)$ 关于 y 对于 $t\in[a,b]$ 是一致连续的.

这时，如果

$$b-a\leqslant\frac{\delta}{M},$$

则方程组(3)至少存在一个满足初始条件(5)的解.

证. 沿用定理 1 的记号并设 $P=P_{y_0}$. 容易看出 P 是从 Ω 到自身的算子. 事实上，在定理 1 证明这一点时仅仅用到不等式 $M(b-a)\leqslant\delta$，这个不等式的成立在本定理中是假定了的.

同样不难验证 P 是连续算子. 事实上，设 $x_n\to x_0$; $x_n\in\Omega$; $z_n=$

$P(x_n)(n=0,1,\cdots)$. 由于函数 $\varphi(y,t)$连续，故对任意的 $\varepsilon>0$ 可找到 $\eta>0$ 使得当$|y-y'|<\eta$ 时，

$$|\varphi(y,t)-\varphi(y',t)|<\varepsilon \qquad (t\in[a,b]). \tag{11}$$

因为$\|x_n-x_0\|\to0$，则当 n 充分大时($n\geqslant n_0$时)有$\|x_n-x_0\|<\eta$，所以更有

$$|x_n(t)-x_0(t)|<\eta \qquad (t\in[a,b]).$$

注意到(11)，所以当 $n\geqslant n_0$时就有

$$|\varphi(x_n(t),t)-\varphi(x_0(t),t)|<\varepsilon \qquad (t\in[a,b]).$$

因而当 $n\geqslant n_0$时

$$\begin{aligned}\|z_n-z_0\|&=\max_{t\in[a,b]}|z_n(t)-z_0(t)|\\&\leqslant\max_{t\in[a,b]}\int_a^t|\varphi(x_n(\tau),\tau)-\varphi(x_0(\tau),\tau)|d\tau\\&\leqslant\varepsilon(b-a).\end{aligned}$$

因此 $z_n=P(x_n)\to z_0=P(x_0)$.

由于 Ω 显然是凸闭集，所以要应用定理 3.2 只要证明集合 $P(\Omega)$是相对紧的就行了. 一旦证得了 $P(\Omega)$的相对紧性，定理的证明也就完成了.

不难将空间 $\mathbf{C}$ 中集合的紧性准则(I.5.2) 移植到空间 $\mathbf{C}_n$ 中. 因此，就应当证明函数族 $P(\Omega)$ 是同等连续的(其一致有界性从 $P(\Omega)\subset\Omega$ 推知，而 Ω 是有界集). 但这乃是下述不等式

$$|z(t')-z(t)|=\left|\int_t^{t'}\varphi(x(\tau),\tau)d\tau\right|\leqslant M|t'-t|$$

$$(x\in\Omega,\ z=P(x))$$

的推论.

定理得证.

4.4. 现在来考察积分方程

$$x(s)=\lambda\int_a^b\varphi(s,t,x(t))dt. \tag{12}$$

应用定理 1.1 可以证明：对于充分小的 λ，方程(12)具有唯一解.

定理 3. 设函数 $\varphi(s,t,u)$ 在平行六面体

$$a \leqslant s,\ t \leqslant b;\ |u| \leqslant h$$

上有定义而且连续，并且对于每个固定的 s 和 t 满足关于 u 的 Lipschitz 条件：

$$|\varphi(s,t,u)-\varphi(s,t,u')| \leqslant K|u-u'|,$$

其中 K 不依赖于 s 和 t.

这时，如果记

$$M=\max|\varphi(s,t,u)| \qquad (a \leqslant s,\ t \leqslant b,\ |u| \leqslant h),$$

则当 λ 满足条件

$$|\lambda| < \min\left[\frac{h}{M(b-a)}, \frac{1}{K(b-a)}\right] \tag{13}$$

时，方程(12)有唯一连续解.

证. 在定理 1.1 中令 $\mathbf{X}=\mathbf{C}[a,b]$ 并取所有使得 $\|x\| \leqslant h$ 的 $x \in \mathbf{C}$ 的全体为 Ω. 算子 P 定义为

$$z=P(x),\ z(s)=\lambda\int_a^b \varphi(s,t,x(t))dt.$$

条件(13)保证了 $P(x) \in \Omega$. 事实上

$$\|P(x)\| \leqslant |\lambda| M(b-a) < h \qquad (x \in \Omega).$$

其次，

$$\begin{aligned}\|P(x)-P(x')\| &\leqslant |\lambda| \max_{s\in[a,b]}\int_a^b |\varphi(s,t,x(t))-\varphi(s,t,x'(t))|dt \\ &\leqslant |\lambda| K(b-a)\|x-x'\|.\end{aligned}$$

因而，若令 $\alpha=|\lambda|K(b-a)$，则由(13)得到 $\alpha<1$，亦即 P 是压缩算子.

由于方程(12)可写成

$$x=P(x), \tag{14}$$

所以方程(12)的解的存在性和唯一性可从定理1.1推知，因为正如

以上所验证过的，定理 1.1 的所有条件均满足.

注 1. 取任意的连续函数 $x_0(\|x_0\|\leqslant h)$，由 x_0 出发构成函数序列 $\{x_n\}$：

$$x_{n+1}(s)=\lambda\int_a^b \varphi(s, t, x_n(t))dt \qquad (n=0, 1, \cdots).$$

这时由定理 1.1 可推知所构成的序列一致收敛于方程(12)的解.

注 2. 利用定理 1.3 则可以建立方程 (12) 的解对满足条件(13)的 λ 的连续依赖性.

4.5. 利用 Schauder 原理有可能在很多情况下扩大使得方程(12)可解性得到保证的 λ 的界.

假定函数 φ 具有

$$\varphi(s, t, u)=K(s, t)\psi(t, u) \tag{15}$$

的形式，其中 $K(s, t)$ 在矩形 $[a, b; a, b]$ 上连续而 $\psi(t, u)$ 对 $t\in[a, b]$，$|u|\leqslant h$ 连续. 这时，有

定理 4. 如果积分方程(12)中的函数具有(15)的形式并且

$$|\lambda|\leqslant\frac{h}{M_1M_2(b-a)}, \tag{16}$$

其中 $M_1=\max\limits_{a\leqslant s, t\leqslant b}|K(s, t)|$，$M_2=\max\limits_{\substack{a\leqslant t\leqslant b\\ |u|\leqslant h}}|\psi(t, u)|$，则方程(12)有连续解.

证. 仍然沿用以前的记号，象上面那样来证明 P 是从集合 Ω 到自身的算子. 由于集合 Ω 是凸闭集而算子 P 连续（见定理 2 的证明)，所以正如在定理 2 中那样，为了能应用定理 3.2 只须验证 $P(\Omega)$ 是相对紧的就够了. 因为 $P(\Omega)$ 是空间 $\mathbf{C}[a, b]$ 的子集合，所以就应当证明 $P(\Omega)$ 中的函数是一致有界和同等连续的. 一致有界性是显然的，而同等连续性则是不等式

$$|z(s)-z(s')|=|\lambda|\left|\int_a^b[K(s,t)-K(s',t)]\psi(t,x(t))dt\right|\leqslant$$

$$\leqslant |\lambda| M_2 \int_a^b |K(s,t)-K(s',t)|\,dt,$$

$$(z=P(x),\ x\in\Omega)$$

的推论.

将积分算子看成空间 $\mathbf{L}^2[a,b]$中的算子,借助于 Schauder 原理可以建立与定理 4 相类似的结果.

定理 5. 设方程(12)中的函数 $\varphi(s,t,u)$具有(15)的形式,函数$K(s,t)$在矩形$[a,b;a,b]$上可测并满足条件

$$A^2=\sup_{s\in[a,b]}\int_a^b |K(s,t)|^2 dt<\infty,$$

而函数 $\psi(t,u)$在矩形$[a,b;-h,h]$上有定义而且连续.

这时,如果

$$|\lambda|\leqslant\frac{h}{AM_2\sqrt{b-a}}, \tag{17}$$

则方程(12)具有在$[a,b]$上平方可积的解.

证. 象上面一样,根据定理 3.2 来证明本定理. 不过在这里我们要设 $\mathbf{X}=\mathbf{L}^2[a,b]$. 取所有使得

$$|x(t)|\leqslant h \qquad (\text{对几乎所有 } t\in[a,b])$$

的 $x\in\mathbf{L}^2$的全体为集合Ω. 算子P仍保持以前的意义.

我们来验证条件(17)保证了包含关系 $P(\Omega)\subset\Omega$. 事实上,如果 $x\in\Omega$ 而 $z=P(x)$,则对 $s\in[a,b]$按 Буняковский不等式有

$$\begin{aligned}|z(s)|&=|\lambda|\left|\int_a^b K(s,t)\psi(t,x(t))\,dt\right|\\&\leqslant|\lambda|\left[\int_a^b |K(s,t)|^2\,dt\right]^{\frac{1}{2}}\left[\int_a^b|\psi(t,x(t))|^2dt\right]^{\frac{1}{2}}\\&\leqslant|\lambda|\,AM_2\sqrt{b-a}\leqslant h.\end{aligned}$$

现在来证明算子P是连续的. 设 x, $x'\in\Omega$ 而 $z=P(x)$, $z'=P(x')$. 因此就有

$$|z(s)-z'(s)|^2 \leqslant |\lambda|^2 \int_a^b |K(s,t)|^2 dt \int_a^b |\psi(t,x(t)) - \psi(t,x'(t))|^2 dt$$

$$\leqslant |\lambda|^2 A^2 \int_a^b |\psi(t,x(t)) - \psi(t,x'(t))|^2 dt. \tag{18}$$

设给定任意的 $\varepsilon>0$. 根据函数 ψ 的连续性, 可找到 $\eta>0$ 使得当 $|u-u'|<\eta$ 时有

$$|\psi(t,u)-\psi(t,u')|<\varepsilon \qquad (t\in[a,b]).$$

其次将区间$[a,b]$分成 e_1 和 e_2 两个集合, 其中 e_1 表示使得 $|x(t)-x'(t)|<\eta$ 的 t 的集合, e_2 为$[a,b]$上其他点的集合. 若记 $\|x-x'\|=\gamma$, 则

$$\gamma^2=\int_a^b |x(t)-x'(t)|^2 dt \geqslant \int_{e_2} |x(t)-x'(t)|^2 dt \geqslant \eta^2 \operatorname{mes} e_2,$$

因此

$$\operatorname{mes} e_2 \leqslant \frac{\gamma^2}{\eta^2}.$$

我们现在来估计积分 $I=\int_a^b |\psi(t,x(t))-\psi(t,x'(t))|^2 dt$:

$$I=\int_{e_1}+\int_{e_2} |\psi(t,x(t))-\psi(t,x'(t))|^2 dt$$

$$\leqslant \varepsilon^2 \operatorname{mes} e_1 + 4M_2^2 \operatorname{mes} e_2$$

$$\leqslant \varepsilon^2 (b-a) + \frac{4M_2^2\gamma^2}{\eta^2}.$$

利用这个估计式, 由(18)可得

$$\|z-z'\|^2=\int_a^b |z(s)-z'(s)|^2 ds$$

$$\leqslant |\lambda|^2 A^2 (b-a)\left[\varepsilon^2 (b-a)+\frac{4M_2^2\|x-x'\|^2}{\eta^2}\right].$$

因此, 只要 x 充分接近 x', z 就可以任意地接近 z'.

由于Ω显然是凸闭集合，所以剩下就是证明集合$P(\Omega)$的相对紧性了. 为此，根据 Riesz 定理(IX. 1. 3)，就只须验证

$$\int_a^b |z(t+\eta)-z(t)|^2\,dt \xrightarrow[\eta\to 0]{} 0$$

对于 $z\in P(\Omega)$是一致地成立就行了*).

设 $x\in\Omega$, $z=P(\Omega)$. 那么

$$\int_a^b |z(s+\eta)-z(s)|^2\,ds$$

$$=\int_a^b \left|\lambda\int_a^b [K(s+\eta,t)-K(s,t)]\psi(t,x(t))\,dt\right|^2 ds$$

$$\leqslant M_2^2\,|\lambda|^2\int_a^b\int_a^b |K(s+\eta,t)-K(s,t)|^2\,ds\,dt.$$

我们现在来证明最后一个积分趋于零.

如果函数 $K(s,t)$连续，则这是显然的. 而如果$K(s,t)$具有一般形式，则可取连续函数序列$\{K_n(s,t)\}$使得

$$\int_a^b\int_a^b |K(s,t)-K_n(s,t)|^2\,ds\,dt \xrightarrow[n\to\infty]{} 0.$$

我们有

$$\left[\int_a^b\int_a^b |K(s+\eta,t)-K(s,t)|^2\,ds\,dt\right]^{\frac{1}{2}}$$

$$\leqslant\left[\int_a^b\int_a^b |K_n(s+\eta,t)-K_n(s,t)|^2\,ds\,dt\right]^{\frac{1}{2}}$$

$$+\left[\int_a^b\int_a^b |K(s+\eta,t)-K_n(s+\eta,t)|^2\,ds\,dt\right.$$

$$+\left[\int_a^b\int_a^b |K(s,t)-K_n(s,t)|^2\,ds\,dt\right]^{\frac{1}{2}}.$$

选 n 充分大，使上面不等式右端后两个积分小于任意的 $\varepsilon>0$. 然后，取 η 充分小使得第一个积分同样小于 ε. 这样，我们就得到在

---

*) 在 Riesz 定理中还有一条件就是范数的有界性，这一点我们没有提及，因为它是显然成立的.

一般情形下所需要的结果.

定理完全得证.

注1. 如果函数 $\psi(t,u)$定义在带形区域 $a\leqslant t\leqslant b$; $-\infty<u<\infty$上，而且

$$\frac{\psi(t,h)}{h}\xrightarrow[h\to\infty]{}0, \tag{19}$$

则当 h 充分大时，定理 4 的条件(16)和定理 5 的条件(17)将对不论怎样的 λ、M_1、M_2 和 $b-a$ 均成立. 因此，(19)式就保证了对任意的 λ 来说方程(12)的解都存在.

注 2. 定理 4 和定理 5 还可以在对方程中所出现的函数施以某些另外的条件的情形加以证明. 同样不难陈述当在不同于 $\mathbf{C}[a,b]$和 $\mathbf{L}^2[a,b]$的其他泛函空间(例如在 $\mathbf{L}[a,b]$中) 中考察方程(12)时的相应的定理.

关于不动点原理对非线性积分方程解的存在性证明的应用可见Красносельский[1].

Schauder 原理的最重要的应用是由 Schauder 本人在其关于椭圆型微分方程的著作中给出的，而 Leray-Schauder 定理是在关于双曲型方程的著作中给出的(见 Schauder[3]).

§ 5. Kakutani 定理

5.1. 本节所述的 Kakutani 定理(参见 Kakutani 的[2]) 乃是 Schauder 定理(3.1) 的推广. Kakutani 定理在数学经济问题中有很多应用(参见Nikaito). 下面，我们叙述这个定理对博弈论的应用.

为了叙述 Kakutani 定理，我们需要关于多值映射的某些事实. 设$\mathbf{X}$和$\mathbf{Y}$是度量空间，并且用 $2^{\mathbf{Y}}$ 表示集合 $\mathbf{Y}$ 的所有子集合的集合. 我们将所有映射 $f:\mathbf{X}\to 2^{\mathbf{Y}}$ 称为多值映射. 由此易见，对

于每个点 $x\in\mathbf{X}$，都有集合 $\mathbf{Y}$ 的某个子集合 $f(x)$ 与之相对应(当然，集合 $\mathbf{Y}$ 我们在以后总认为是非空的).

我们称如下定义的 $\mathbf{X}\times\mathbf{Y}$ 中的子集 G_f 为多值映射 $f:\mathbf{Y}\to 2^{\mathbf{Y}}$ 的图形:

$$G_f=\{(x,y):y\in f(x)\}.$$

多值映射 $f:\mathbf{X}\to 2^{\mathbf{Y}}$ 称为在点 $x\in\mathbf{X}$ 处是闭的，是指从 $x_n\to x$，$y_n\to y$ 和 $y_n\in f(x_n)$ 可推出 $y\in f(x)$. 如果映射 f 在每个点 $x\in\mathbf{X}$ 处都是闭的，则称映射 f 是闭的. 易见，f 是闭映射的充分必要条件是 f 的图形 G_f 在 $\mathbf{X}\times\mathbf{Y}$ 中是闭的.

多值映射 $f:\mathbf{X}\to 2^{\mathbf{Y}}$ 称作在点 $x\in\mathbf{X}$ 处是上半连续的，乃是指对任意的开集合 $U\supset f(x)$ 均可找到点 x 的邻域 V，使得 $f(V)\subset U$，其中 $f(V)\equiv\bigcup\limits_{z\in V}f(z)\subset U$. 如果在每个点 $x\in\mathbf{X}$ 处映射 f 都是上半连续的，则称映射 f 是上半连续的.

如果 f 是通常的单值映射，则 f 的上半连续性等价于连续性，虽然闭映射可能不是连续映射.

引理 1. 如果集合 $\mathbf{Y}$ 是紧的，则闭映射 f 上半连续.

证. 取 $x\in\mathbf{X}$ 和开集合 $U\supset f(x)$. 考察集合 $V=\{z\in\mathbf{X}:\ f(z)\subset U\}$. 因为 $x\in V$ 并且 $f(V)\subset U$，所以只需证明 V 是开集合，或者 $\mathbf{X}\backslash V=\{z\in\mathbf{X}:\ f(z)\cap(\mathbf{Y}\backslash U)\neq\varnothing\}$ 是闭集. 如果序列 $z_n\to z\in\mathbf{X}$ 而 $z_n\in\mathbf{X}\backslash V$，则可以找到点 $y_n\in f(z_n)\cap(\mathbf{Y}\backslash U)$. 由于 $\mathbf{Y}$ 的紧性，可以找到序列 $\{y_{n_k}\}$ 使得 $y_{n_k}\to y\in\mathbf{Y}\backslash U$. 因为映射 f 是闭的，所以可推知 $y\in f(z)$. 因而 $y\in f(z)\cap(\mathbf{Y}\backslash U)$，由此可知 $z\in X\backslash V$.

5.2. 点 $x^*\in\mathbf{X}$ 称作映射 $f:\mathbf{X}\to 2^{\mathbf{X}}$ 的不动点，乃是指 $x^*\in f(x^*)$. 现在可以来叙述 Kakutani 定理了.

定理 1. 设 $\mathbf{X}$ 是 Banach 空间，K 是 $\mathbf{X}$ 中的非空紧凸集合，而 $f:K\to 2^k$ 是满足条件

1）对于每个点 $x\in K$，集合 $f(x)$ 均是集合 K 的非空凸子集；

2）映射 f 是闭的，

的多值映射.

这时，映射 f 具有不动点.

证. 因为 K 是紧集合，所以按照 Hausdorff 定理，对任意的 $\varepsilon>0$，存在关于 K 的有限的 ε-网 $x_{\varepsilon 1}$，$x_{\varepsilon 2}$，…，$x_{\varepsilon n(\varepsilon)}$. 对任意的 $x\in K$，令

$$\varphi_{\varepsilon i}(x)=\max(\varepsilon-\|x-x_{\varepsilon i}\|,0)\qquad (i=1,2,\cdots,n(\varepsilon)).$$

显然，$\varphi_{\varepsilon i}$ 是 K 上的连续非负函数. 因为 $\{x_{\varepsilon i}\}_{i=1}^{n(\varepsilon)}$ 是 ε-网，所以对任意的 $x\in K$，至少有某个 i 使得 $\|x-x_{\varepsilon i}\|<\varepsilon$. 因此，对这个 i 就有 $\varphi_{\varepsilon i}(x)>0$. 这表明下述定义

$$w_{\varepsilon i}(x)=\varphi_{\varepsilon i}(x)\Big/\sum_{j=1}^{n(\varepsilon)}\varphi_{\varepsilon j}(x)\qquad (i=1,2,\cdots,n(\varepsilon))\qquad (*)$$

的正确性（也就是说，鉴于对上面所讲的 i 有 $\varphi_{\varepsilon i}(x)>0$，所以可以按 $(*)$ 来定义 $w_{\varepsilon i}(x)$——译注）.

取定任意的 n 个点 $y_{\varepsilon i}\in f(x_{\varepsilon i})\ (i=1,\cdots,n(\varepsilon))$ 并按公式

$$f_\varepsilon(x)=\sum_{i=1}^{n(\varepsilon)}w_{\varepsilon i}(x)y_{\varepsilon i}$$

确定从 K 到 K 的单值连续映射 f_ε.

从条件 $y_{\varepsilon i}\in K$，$w_{\varepsilon i}(x)\geqslant 0$，$\sum w_{\varepsilon i}(x)=1$ 以及由集合 K 的凸性可推知 $f_\varepsilon(x)\in K$. 这样，对任意的 $\varepsilon>0$ 就有单值连续映射 f_ε: $K\to K$. 根据 Schauder 定理（见 3.1），存在映射 f_ε 的不动点 x_ε，即 $f_\varepsilon(x_\varepsilon)=x_\varepsilon$.

由于集合 K 的紧性，可求得正数序列 $\{\varepsilon_n\}$ 和点 $x^*\in K$ 使得满足如下的条件：

---

*）定理 1 的证明是依照 Никайдо 的方法得到的.

a) $\lim\limits_{n\to\infty}\varepsilon_n=0$; b) $x_{\varepsilon_n}\to x^*$; c) $f_{\varepsilon_n}(x_{\varepsilon_n})=x_n$.

我们现在来证明 x^* 就是要求的映射 f 的不动点. 命 $U_\delta=f(x^*)+B_\delta$, 其中 $B_\delta=\{y:\|y\|<\delta\}$, $\delta>0$. 若能证明对任意的 $\delta>0$ 均有 $x^*\in U_\delta$, 则由于 $f(x^*)$ 的闭性, 我们就得到了所要证明的结果: $x^*\in f(x^*)$. (注意, $f(x^*)$ 的闭性从映射 f 的闭性推知).

现在来证明: 对任意的 $\delta>0$ 均有 $x^*\in U_\delta$. 显然, U_δ 是凸开集而 $f(x^*)\subset U_\delta$. 根据引理 1, 对于 x^* 和 U_δ 可以找到球 $K_\varepsilon=\{x:\|x-x^*\|<\varepsilon\}$ 使得 $f(K_\varepsilon)\subset U_\delta$. 由于条件 a) 和 b), 存在正数 N 使得当 $n\geqslant N$ 时有 $\varepsilon_n<\varepsilon/2$ 并且 $x_{\varepsilon_n}\in K_{\varepsilon/2}$. 如果 $w_{\varepsilon_n i}(x_{\varepsilon_n})>0$ 则

$$\|x_{\varepsilon_n i}-x_{\varepsilon_n}\|<\varepsilon_n<\varepsilon/2,$$

从而

$$\|x_{\varepsilon_n i}-x^*\|\leqslant\|x_{\varepsilon_n i}-x_{\varepsilon_n}\|+\|x_{\varepsilon_n}-x^*\|<\varepsilon/2+\varepsilon/2=\varepsilon.$$

因而, 当 $n>N$ 时, 对所有使得 $w_{\varepsilon_n i}(x_{\varepsilon_n})>0$ 的 i 均有 $x_{\varepsilon_n i}\in K_\varepsilon$. 对这样的 i, 我们就得到

$$y_{\varepsilon_n i}\in f(x_{\varepsilon_n i})\subset f(K_\varepsilon)\subset U_\delta. \tag{1}$$

从条件 c) 可以推知

$$x_{\varepsilon_n}=\sum w_{\varepsilon_n i}(x_{\varepsilon_n})y_{\varepsilon_n i}. \tag{2}$$

由(1)和(2)我们就得到结果: 当 $n\geqslant N$ 时, 点 x_{ε_n} 是所有位于 U_δ 中的点 $y_{\varepsilon_n i}$ 的凸组合, 再注意到 U_δ 的凸性, 于是可知 $x_{\varepsilon_n}\in U_\delta$. 令 $n\to\infty$ 而取极限, 得到 $x^*\in\bar{U}_\delta\subset U_{2\delta}$. 由此, 正如早已指出的那样, 可推知 $x^*\in f(x^*)$. 这就最终完成 Kakutani 定理的证明.

5.3. 我们现在来给出 Kakutani 定理对博弈论的一个应用.

考虑具有零和的两个人(赌者 I 和赌者 II)的博弈, 也就是说博弈中的一个参加者的赢得量只能等于另一个参加者的输掉量. 我们将赌者在一次假定性的博弈过程中所可能采取的全部行动的所有选择叫赌者的策略. 赌者 I 的所有策略的集合 (亦即策略空间)记为 X, 赌者 II 的所有策略的集合记为 Y.

定义在 $X\times Y$ 上的实函数 $K(x,y)$ 称为利益函数（Функция выигрыша），并且若赌者 I 选择策略 $x\in X$ 而赌者 II 选择策略 $y\in Y$，则数 $K(x,y)$ 就被解释成是赌者 I 的利益．而 $-K(x,y)$ 是赌者 II 的利益．三元组（X，Y，K）称为博弈，其中 X 和 Y 是如上所述的两个集合而 K 是定义在 $X\times Y$ 上的函数．

当赌者 I 选择自己的策略 $x\in X$ 时，他根本不知道赌者 II 的策略．因此赌者 I 应该设法使他的对手选择的是最坏的策略，亦即赌者 I 的利益是 $\inf\limits_{y\in Y}K(x,y)$．这样，赌者 I 自然要选择策略 $x_0\in X$ 使得

$$\inf_{y\in Y}K(x_0,y)=\max_{x\in X}\inf_{y\in Y}K(x,y). \tag{3}$$

由于赌者 II 的利益是 $-K(x,y)$，所以从(3)可知赌者 II 应该找这样的策略 $y_0\in Y$，使得

$$\inf_{x\in X}\{-K(x,y_0)\}=\max_{x\in X}\inf_{y\in Y}[-K(x,y)],$$

亦即使得

$$\sup_{x\in X}K(x,y_0)=\min_{y\in Y}\sup_{x\in X}K(x,y). \tag{4}$$

比较(3)和(4)可看出，要得到存在满足(3)和(4)的策略 $x_0\in X$ 和 $y_0\in Y$，其充分必要条件是对所有的 $x\in X$ 和 $y\in Y$ 要满足不等式

$$K(x,y_0)\leqslant K(x_0,y_0)\leqslant K(x_0,y), \tag{5}$$

而(5)式成立的充分必要条件是满足等式

$$\min_{y\in Y}\sup_{x\in X}K(x,y)=\max_{x\in X}\inf_{y\in Y}K(x,y). \tag{6}$$

使等式(6)成立的值称为博弈的值．满足不等式(5)的策略偶 x_0 和 y_0 称为博弈的解，策略 x_0 和 y_0 本身则称为最佳策略*)．为了证明博弈的解的存在性需要证明一个关于极小极大的定理，这一定理在于对给定的博弈来验证等式（6）．我们现在借助于 Ka-

*) 详细的论述可参见 Карлин．本节中部分地引用该书的内容．

kutani定理来证明这类定理中的一个最一般的定理. 首先叙述下面的引理, 这个引理的证明建议读者自己去做.

引理 2. 1) 对任何实函数$K(x, y)$, 均有

$$\inf_{y\in Y}\sup_{x\in X}K(x, y)\geqslant\sup_{x\in X}\inf_{y\in Y}K(x, y). \tag{7}$$

2) 如果X和Y是度量紧的而函数$K(x, y)$在$X\times Y$上连续, 则函数$\varphi(x)=\max\limits_{y\in Y}K(x, y)$和函数$\psi(x)=\min\limits_{y\in Y}K(x, y)$在$X$上连续. 此外, 存在$\min\limits_{y\in Y}\max\limits_{x\in X}K(x, y)$和$\max\limits_{x\in X}\min\limits_{y\in Y}K(x, y)$.

定理 2. 设以下诸条件满足:

a) X是Banach空间$\mathbf{X}$中的凸紧集;

b) Y是Banach空间$\mathbf{Y}$中的凸紧集;

c) 函数$K(x, y)$在$X\times Y$上连续, 对于每个$x\in X$, $K(x, y)$是y的凸函数, 对于每个$y\in Y$, $K(x, y)$是x的凹函数, 也就是说

$$K(x, \lambda y_1+(1-\lambda)y_2)\leqslant\lambda K(x, y_1)+(1-\lambda)K(x_2, y),$$

$$K(\lambda x_1+(1-\lambda)x_2, y)\geqslant\lambda K(x_1, y)+(1-\lambda)K(x_2, y),$$

其中$x, x_1, x_2\in X$; $y, y_1, y_2\in Y$; $0\leqslant\lambda\leqslant1$.

这时, 博弈(X, Y, K)有解, 而且

$$\min_{y\in Y}\max_{x\in X}K(x, y)=\max_{x\in X}\min_{y\in Y}K(x, y). \tag{8}$$

证. 从引理2的2)可知, 使得博弈有解的条件(6)和(8)事实上是一回事. 因此, 我们将证明(8)式的成立.

对于任何$x\in X$, 令

$$B_x=\{y\in Y: K(x, y)=\min_{z\in Y}K(x, z)\}.$$

因为集合Y是紧的, 故$B_x\neq\varnothing$。易见B_x是闭的. 今证B_x还是凸的. 如果$y_1, y_2\in B_x$, $0\leqslant\lambda\leqslant1$, 则

$$K(x, \lambda y_1+(1-\lambda)y_2)\leqslant\lambda K(x, y_1)+(1-\lambda)K(x, y_2)$$

$$=\lambda\min_{z\in Y}K(x, z)+(1-\lambda)\min_{z\in Y}K(x, z)=\min_{z\in Y}K(x, z). \tag{9}$$

另一方面, 因为Y是凸的, 所以$\lambda y_1+(1-\lambda)y_2\in Y$, 由此推知

$$K(x, \lambda y_1+(1-\lambda) y_2) \geqslant \min_{z \in Y} K(x, z). \tag{10}$$

由(9)和(10)可得到

$$K(x, \lambda y_1+(1-\lambda) y_2)=\min_{z \in Y} K(x, z).$$

由此可知 $\lambda y_1+(1-\lambda) y_2 \in B_x$, 也就是说集合 B_x 是凸的.

类似地可以证明, 对任何 $y \in Y$, 集合

$$A_y=\{x \in X: K(x, y)=\max_{w \in X}(w, y)\}$$

也是非空的闭凸集.

考察 Banach 空间 $\mathbf{X} \times \mathbf{Y}$ 中的凸集合 $C=X \times Y$. 根据Тихонов 定理(I. 2. 7), 集合 C 在空间 $\mathbf{X} \times \mathbf{Y}$ 中是紧的. 显然, $A_y \times B_x$ $(x \in X, y \in Y)$ 是 C 中的非空凸闭子集, 我们现在来考察将 $(x, y) \in C$ 映射到集合 $A_y \times B_x$ 内去的多值映射 f: $C \rightarrow 2^C$. 为了能应用 Kakutani 定理, 只需证明映射 f 是闭的就可以了.

设在集合 C 中有 $(x_n, y_n) \rightarrow(x, y)$, $(u_n, v_n) \rightarrow(u, v)$ 而且 $(u_n, v_n) \in f(x_n, y_n)$. 我们来证明

$$(u, v) \in f(x, y)=A_y \times B_x.$$

因为 $u_n \in A_{y_n}$, $v_n \in B_{x_n}$, 所以

$$K(u_n, y_n)=\max_{w \in X} K(w, y_n),$$

$$K(x_n, v_n)=\min_{z \in Y} K(x_n, z).$$

因为 $u_n \rightarrow u$, 所以由引理 2 的 2)可知

$$K(u, y)=\lim_{n \rightarrow \infty} K(u_n, y_n)=\lim_{n \rightarrow \infty} \max_{w \in X} K(w, y_n)$$

$$=\max_{w \in X} K(w, y).$$

因而证得 $u \in A_y$. 类似地可证 $v \in B_x$.

这就证明了映射 f 是闭的. 由于存在着点 $(x_0, y_0) \in C$ 使得 $(x_0, y_0) \in f(x_0, y_0)=A_{y_0} \times B_{x_0}$, 应用 Kakutani 定理可知

$$K(x_0, y_0)=\max_{x \in X} K(x, y_0), \quad K(x_0, y_0)=\min_{y \in Y} K(x_0, y).$$

由此可得

$$\min_{y\in Y}\max_{x\in X}K(x,y)\leqslant K(x_0,y_0)\leqslant\max_{x\in X}\min_{y\in Y}K(x,y).$$

将上不等式与(7)比较即得到欲证的(8)。

定理完全得证.

第十七章　非线性算子的微分

非线性算子的进一步研究可以通过建立线性算子和非线性算子之间的联系这个途径来实现，讲详细点，就是借助于由线性算子去局部地逼近非线性算子的途径来实现.

为此，本章里对非线性算子微分学这个工具给予必要的讨论(这种讨论完全是为了下一章的应用).

§1. 一 阶 导 数

1.1. 设P是从一个 Banach 空间$\mathbf{X}$的闭集合Ω到另一个 Banach 空间$\mathbf{Y}$的集合Δ内的算子. 取固定的一个元素$x_0\in\Omega$，并假定存在这样一个连续线性算子$U\in B(\mathbf{X},\mathbf{Y})$*)，使得对任何$x\in\mathbf{X}$

$$\lim_{t\to0}\frac{P(x_0+tx)-P(x_0)}{t}=U(x).\tag{1}$$

在这种情况下，就说线性算子U是算子P在点x_0处的导数，并写为

$$U=P'(x_0).$$

如上所定义的导数常称为 Gateaux 导数或弱导数，而元素$U(x)$称做 Gateaux 微分.

用$\overline{K}$表示使得$\|x\|=1$的所有$x\in\mathbf{X}$的集合. 如果极限关系(1)对$x\in\overline{K}$一致地成立，则称算子P在点x_0处可微，这时导数$P'(x_0)$叫做 Fréchet 导数或强导数**).

换句话说，算子P在点x_0处的可微性表明存在这样的算子

---

*) 记号$B(\mathbf{X},\mathbf{Y})$表示从$\mathbf{X}$到$\mathbf{Y}$的连续线性算子全体组成的空间.

**) 关于 Fréchet 导数的定义在 Fréchet[3]中给出，弱导数的定义在 Gateaux[1]中给出.

$U\in B(\mathbf{X},\mathbf{Y})$，使得对任何 $\varepsilon>0$，均可找到 $\delta>0$，从 $\|\Delta x\|<\delta$ 可推出

$$\|P(x_0+\Delta x)-P(x_0)-U(\Delta x)\|\leqslant\varepsilon\|\Delta x\| \qquad (\Delta x\in\mathbf{X}). \qquad (2)$$

这个结论的简单的证明建议读者去做.

1.2. 下面我们指出导数的若干最简单的性质.

I. 如果算子 P 在点 x_0 处可微，则它在这点连续（只假定导数存在还不够）.

这从(2)式可直接推得.

II. 设 $P=\alpha_1P_1+\alpha_2P_2$. 如果存在导数 $P_1'(x_0)$ 和 $P_2'(x_0)$，则也存在 $P'(x_0)$，而且

$$P'(x_0)=\alpha_1P_1'(x_0)+\alpha_2P_2'(x_0).$$

这时，若 P_1 和 P_2 在 x_0 点可微，则 P 在 x_0 点也可微.

这些结论可从定义直接推知.

III. 如果 $P=U\in B(\mathbf{X},\mathbf{Y})$，则 P 在每个点 $x_0\in\mathbf{X}$ 都可微，而且

$$P'(x_0)=U.$$

事实上，

$$\frac{P(x_0+tx)-P(x_0)}{t}=U(x).$$

IV. 设 P 是从 $\Omega\subset\mathbf{X}$ 到开集合 $\Delta\subset\mathbf{Y}$ 的算子，Q 是映 Δ 到某个 Banach 空间 $\mathbf{Z}$ 的算子. 令 $R=QP$，并设 Q 在点 $y_0=P(x_0)$ 处可微（$x_0\in\Omega$），而 P 在点 x_0 处具有导数，那么，算子 R 在点 x_0 处具有导数，并且

$$R'(x_0)=Q'(P(x_0))P'(x_0)=Q'(y_0)P'(x_0).$$

事实上，写 $U=P'(x_0)$，$V=Q'(y_0)$，并取任意的 $x\in\mathbf{X}$. 令 $\Delta y=P(x_0+tx)-P(x_0)$，则

$$\frac{R(x_0+tx)-R(x_0)}{t}=\frac{QP(x_0+tx)-QP(x_0)}{t}$$

$$=\frac{Q(y_0+\Delta y)-Q(y_0)}{t}=\frac{V(\Delta y)+\zeta(\Delta y)}{t}=$$

$$=V\left(\frac{P(x_0+tx)-P(x_0)}{t}\right)+\frac{\zeta(\Delta y)}{\|\Delta y\|}\left\|\frac{P(x_0+tx)-P(x_0)}{t}\right\|, \qquad (3)$$

其中

$$\zeta(\Delta y)\doteq Q(y_0+\Delta y)-Q(y_0)-V(\Delta y),\quad \|\zeta(\Delta y)\|=o(\|\Delta y\|).$$

在(3)中当 $t\to 0$时，第一项的极限为

$$[VP'(x_0)](x)=VU(x),$$

而第二项的极限趋于零(因为$\frac{\zeta(\Delta y)}{\|\Delta y\|}\to 0$ 而范数因子是有界的).

于是，
$$\lim_{t\to 0}\frac{R(x_0+tx)-R(x_0)}{t}=VU(x),$$

这正是所要证的.

注. 如果P在 x_0 点可微，则R同样可微.

V. 如果在IV的假定条件下，算子 $Q=V$ 连续且线性，则

$$R'(x_0)=VP'(x_0).$$

1.3. 我们现在指出一个不等式，普通实函数的有限增量公式正是这个不等式的特例.

设 x_0 和 x 都属于Ω；还假定连接 x_0 和 x 的直线段$[x_0, x]$上的所有点也属于Ω，又设在$[x_0, x]$上的每一点都存在算子P的导数. 写 $\Delta x=x-x_0$，取任意的泛函 $g\in \mathbf{Y}^*$，命

$$\varphi(t)=g(P(x_0+t\Delta x)). \qquad (4)$$

不难看到，实函数 φ 在$[0,1]$上具有导数. 事实上，

$$\frac{\varphi(t+\Delta t)-\varphi(t)}{\Delta t}=g\left(\frac{P(x_0+t\Delta x+\Delta t\Delta x)-P(x_0+t\Delta x)}{\Delta t}\right). \qquad (5)$$

当 $\Delta t\to 0$ 时，泛函 g 号下的表达式具有极限 $P'(x_0+t\Delta x)(\Delta x)$，因而，由泛函 g 的连续性，(5)的右端具有极限

$$\varphi'(t)=g(P'(x_0+t\Delta x)(\Delta x)). \qquad (6)$$

写出对函数 φ 的有限增量公式

$$\varphi(1)-\varphi(0)=\varphi'(\theta) \qquad (0<\theta<1).$$

注意到(4)与(6)，得到

$$g(P(x_0+\Delta x)-P(x_0))=g(P'(x_0+\theta\Delta x)(\Delta x)),$$

因而，

$$|g(P(x_0+\Delta x)-P(x_0))|\leqslant\|g\|\sup_{0<\theta<1}\|P'(x_0+\theta\Delta x)\|\|\Delta x\|. \tag{7}$$

选择非零泛函 g 使得

$$g(P(x_0+\Delta x)-P(x_0))=\|g\|\|P(x_0+\Delta x)-P(x_0)\|.$$

这时，由(7)得到

$$\|P(x)-P(x_0)\|\leqslant\sup_{0<\theta<1}\|P'(x_0+\theta\Delta x)\|\|\Delta x\| \quad (\Delta x=x-x_0). \tag{8}$$

对算子 $P-U$（其中 $U\in B(\mathbf{X},\mathbf{Y})$）应用(8)式并利用 1.2 的 II 和 III，我们就得到不等式

$$\|P(x)-P(x_0)-U(\Delta x)\|\leqslant\sup_{0<\theta<1}\|P'(x_0+\theta\Delta x)-U\|\|\Delta x\|.$$

特别，令 $U=P'(x_0)$，我们就得到

$$\|P(x_0+\Delta x)-P(x_0)-P'(x_0)\Delta x\|$$
$$\leqslant\|\Delta x\|\sup_{0<\theta<1}\|P'(x_0+\theta\Delta x)-P'(x_0)\|. \tag{9}$$

我们称(8)为有限增量公式，而称（9）为带余项的有限增量公式.

1.4. 我们指出有限增量公式的一些应用.

假定算子 P 的导数在某个开集合 $\Omega_0\subset\Omega$ 上的每一点都存在. 因此，在 Ω_0 中就定义了一个把元素 $x\in\Omega_0$ 和元素 $P'(x)\in B(\mathbf{X},\mathbf{Y})$ 对应起来的算子 P'. 今证：如果算子 P' 在点 $x_0\in\Omega_0$ 处连续，则 P 在这点可微.

事实上，由于假定的算子 P' 的连续性，对不管怎样的 $\varepsilon>0$，总可以找到 $\delta>0$，使得只要 $\|x-x_0\|<\delta$，就有

$$\|P'(x)-P'(x_0)\|<\varepsilon,$$

也就是说，如果$\|\Delta x\|<\delta$，则$\sup\limits_{0<\theta<1}\|P'(x_0+\theta\Delta x)-P'(x_0)\|\leqslant\varepsilon$. 把这个结果应用于(9)，即得到不等式

$$\|P(x_0+\Delta x)-P(x_0)-P'(x_0)(\Delta x)\|\leqslant\varepsilon\|\Delta x\|,$$

这表明算子P在点x_0处的可微性.

有限增量公式还可以用以证明算子P的不动点的存在性.

定理 1. 设算子P映集合$\Omega\subset\mathbf{X}$到$\mathbf{X}$，而且在凸闭集合$\Omega_0\subset\Omega$的每一点存在导数. 这时，如果

1) $P(\Omega_0)\subset\Omega_0$;

2) $\sup\limits_{x\in\Omega_0}\|P'(x)\|=\alpha<1$,

则在Ω_0中存在算子P的唯一的不动点.

证. 只须验证在所考察的情况下定理 XVI.1.1 的条件满足，也就是只要去说明P是压缩算子.

设$x_1, x_2\in\Omega_0$. 按有限增量公式，有

$$\|P(x_2)-P(x_1)\|\leqslant\|x_2-x_1\|\sup_{0<\theta<1}\|P'(x_1+\theta(x_2-x_1))\|$$

$$\leqslant\alpha\|x_2-x_1\|,$$

这就是需要证明的.

注. 如果假定算子P'在Ω_0中连续，则 2) 也是使得P成为压缩算子的必要条件.

事实上，如果

$$\sup_{x\in\Omega_0}\|P'(x_0)\|\geqslant1,$$

则可找到数列$\{\alpha_n\}$ $(\alpha_n\to1)$和序列$\{x_n\}\subset\Omega_0$，使得

$$\|P'(x_n)\|>\alpha_n\qquad(n=1,2,\cdots);$$

同时可认为点x_n是Ω_0的内点.

对每个$n=1,2,\cdots$，按导数的定义，可找到元素x'_n使得对充分小的t有

$$\left\|\frac{P(x_n+tx'_n)-P(x_n)}{t}\right\|>\alpha_n\|x'_n\|\qquad(|t|\leqslant t_n;n=1,2,\cdots).$$

令 $y_n=x_n+t_n x'_n$，我们即得到这样一对元素 x_n 和 y_n，使得

$$\frac{\|P(y_n)-P(x_n)\|}{\|y_n-x_n\|}>\alpha_n, \tag{10}$$

也就是说 P 不是压缩算子.

1.5. 我们来说明：当空间 $\mathbf{X}$ 和 $\mathbf{Y}$ 中一个或两个都是有限维空间时上面所引进的概念的意义.

先设 $\mathbf{X}$ 是一维实空间，即它的元素是由实数构成.

设 $U\in B(\mathbf{X},\mathbf{Y})$. 不难看到这时算子 U 具有如下形式

$$U(t)=ty_0 \qquad (t\in\mathbf{X}), \tag{11}$$

其中 y_0 是 $\mathbf{Y}$ 中某元素. 此外，由于 $\|U(t)\|=|t|\|y_0\|$，则

$$\|U\|=\|y_0\|. \tag{12}$$

反之，对不管怎样的 $y_0\in\mathbf{Y}$，(11)式给出了算子 $U\in B(\mathbf{X},\mathbf{Y})$. 易见，空间 $\mathbf{Y}$ 和 $B(\mathbf{X},\mathbf{Y})$ 中的元素之间的对应关系是一对一的，线性的，由(12)知还是保持范数的. 因此，可以认为 $B(\mathbf{X},\mathbf{Y})=\mathbf{Y}$.

现在来考察从 $\Omega\subset\mathbf{X}$ 到 $\mathbf{Y}$ 的算子 F. 因此，F 就是以实数为自变量而取值于空间 $\mathbf{Y}$ 的函数.

导数 $F'(t_0)=y_0\in\mathbf{Y}$ 存在就表示对任何 $t\in\mathbf{X}$

$$\lim_{\tau\to 0}\frac{F(t_0+\tau t)-F(t_0)}{\tau}=ty_0. \tag{13}$$

记 $\Delta t=\tau t$，可把(13)式改写为

$$\lim_{\Delta t\to 0}\frac{F(t_0+\Delta t)-F(t_0)}{\Delta t}=y_0.$$

于是，在我们所考察的情况下，$F'(t_0)$ 的定义和实变量实函数的普通的导数定义无任何不同. 需要指出的是，由于 $F'(t_0)\in\mathbf{Y}$，所以算子 F' 和 F 一样是映某个数集合到 $\mathbf{Y}$ 的算子.

正象在实函数中的情形一样，可以证明：导数 $F'(t_0)$ 的存在就保证了算子 F 在点 t_0 处可微.

如果 $\Omega=[a,b]$，则集合 $F(\Omega)$ 自然视为空间 $\mathbf{Y}$ 中的曲线，因

而，这时元素 $F'(t_0)$ 就给出了该曲线 $F(\Omega)$ 在点 $y_0=F(t_0)$ 处的切线方向.

1.6. 现在设 $\mathbf{X}$ 和 $\mathbf{Y}$ 是有限维空间. 分别以 m 和 n 表示 $\mathbf{X}$ 和 $\mathbf{Y}$ 的维数. 正如在 V.2.8 中所指出的，任一算子 $U\in B(\mathbf{X},\mathbf{Y})$ 由矩阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \tag{14}$$

按下述公式

$$y=U(x) \qquad (y=(\eta_1,\eta_2,\cdots,\eta_n)\in\mathbf{Y},\ x=(\xi_1,\xi_2,\cdots,\xi_m)\in\mathbf{X}),$$

$$\eta_i=\sum_{k=1}^{m}a_{ik}\xi_k \qquad (i=1,2,\cdots,n). \tag{15}$$

来确定.

考察从 $\Omega\subset\mathbf{X}$ 到 $\mathbf{Y}$ 的某个算子 P. 算子 P 由 m 个变量的 n 个数函数 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 定义如下：如果

$$y=P(x) \qquad (y=(\eta_1,\eta_2,\cdots,\eta_n)\in\mathbf{Y},\ x=(\xi_1,\xi_2,\cdots,\xi_m)\in\Omega),$$

则

$$\eta_i=\varphi_i(\xi_1,\xi_2,\cdots,\xi_m) \qquad (i=1,2,\cdots,n). \tag{16}$$

假定存在导数 $P'(x_0)=U(x_0=(\xi_1^{(0)},\xi_2^{(0)},\cdots,\xi_m^{(0)}))$. 设 U 由矩阵(14)决定. 把 $P'(x_0)=U$ 改写为

$$\lim_{t\to 0}\frac{P(x_0+tx)-P(x_0)}{t}=U(x)$$

并注意到(15)和(16)，我们就得到 n 个关系式

$$\lim_{t\to 0}\frac{\varphi_i(\xi_1^{(0)}+t\xi_1,\xi_2^{(0)}+t\xi_2,\cdots,\xi_m^{(0)}+t\xi_m)-\varphi_i(\xi_1^{(0)},\cdots,\xi_m^{(0)})}{t}$$

$$=\sum_{k=1}^{m}a_{ik}\xi_k \qquad (i=1,2,\cdots,n). \tag{17}$$

因为这些等式应对所有的 $x\in\mathbf{X}$ 成立，如果我们逐次地取这样的元素 x：第一个元素除去第一个坐标等于 1 外其余坐标都等于零，第二个元素除去第二个坐标等于 1 外其余的坐标均等于零，等等，则我们就得到函数 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 关于 $\xi_1,\xi_2,\cdots,\xi_m$ 的偏导数，即

$$\frac{\partial\varphi_i(\xi_1^{(0)},\cdots,\xi_m^{(0)})}{\partial\xi_k}=a_{ik}\qquad(i=1,2,\cdots,n;k=1,2,\cdots,m).$$

因此，导数 $P'(x_0)$ 就是由函数 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 的偏导数为元素构成的矩阵所定义的线性算子.

但需指出，函数 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 的偏导数的存在并不能保证导数 $P'(x_0)$ 的存在，这可由下例说明：设 $n=1$, $m=2$.

$$\varphi_1(\xi_1,\xi_2)=\frac{\xi_1\xi_2}{(\xi_1^2+\xi_2^2)^2},\quad \varphi_1(0,0)=0;x_0=(0,0).$$

易见，$\dfrac{\partial\varphi_1(0,0)}{\partial\xi_1}=\dfrac{\partial\varphi_1(0,0)}{\partial\xi_2}=0$. 因此，如果导数 $P'(x_0)$ 存在，这个导数就会是零算子，因而，(17)式就给出

$$\lim_{t\to0}\frac{\varphi_1(t\xi_1,t\xi_2)}{t}=0.$$

然而，只要 $\xi_1\neq0,\xi_2\neq0$，这个极限趋于无穷大.

我们建议读者证明：算子 P 可微的充分必要条件是函数 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 可微. 但正如由初等分析的一个著名事实可推知的由导数的存在不能推出算子 P 的可微性.

1.7. 设 F 是取值于 Banach 空间 $\mathbf{X}$ 的实变量函数. 如果 F 定义在 $[a,b]$ 上，则论及 F 的积分是有意义的. 我们在 $[a,b]$ 上构造如下的积分和的极限：

$$\sum_{k=0}^{n-1}F(\tau_k)(t_{k+1}-t_k)$$

$(a=t_0<t_1<\cdots<t_n=b;\tau_k\in[t_k,t_{k+1}],k=0,1,\cdots,n-1)$

$$（\text{当}\ \lambda=\max_k[t_{k+1}-t_k]\to 0\ \text{时}）.$$

当上述极限存在时，就称为函数 F 的积分并写成 $\int_a^b F(t)dt$. 正象对实值函数一样，可以证明闭区间上的连续函数是一致连续的，因而是可积的. 也就是说，对于连续函数，它的积分存在. 普通 Riemann 积分的性质及证明都可移植到如上所定义的抽象积分上去. 我们只举其中的三个性质.

I. 如果 $U\in B(\mathbf{X},\mathbf{Y})$，则

$$\int_a^b U(F(t))\,dt=U\left(\int_a^b F(t)\,dt\right).$$

II. 如果 $F(t)=\varphi(t)x_0$ $(t\in[a,b])$ 而 $x_0\in\mathbf{X}$ 是固定元素，$\varphi(t)$ 是实的可积函数，则

$$\int_a^b F(t)dt=x_0\int_a^b \varphi(t)\,dt.$$

III. $$\left\|\int_a^b F(t)\,dt\right\|\leqslant\int_a^b \|F(t)\|\,dt.$$

我们现在再回到对非线性算子的研究.

设 R 是定义在区间 $[x_0,x_0+\Delta x]$ 上而取值于空间 $B(\mathbf{X},\mathbf{Y})$ 中的某个算子. 由定义，我们设

$$\int_{x_0}^{x_0+\Delta x} R(x)\,dx=\int_0^1 R(x_0+t\Delta x)\Delta x\,dt$$

$$=\lim\sum_{k=0}^{n-1}R(x_0+\tau_k\Delta x)\Delta x(t_{k+1}-t_k).$$

显然，如果 R 是连续的，则积分存在且该积分为 $\mathbf{Y}$ 中的元素. 这种类型的积分最先是由 М. К. Гаврин[1]引入的.

特别，如果 $R=P'$，P 是由 $\Omega\subset\mathbf{X}$ 到 $\mathbf{Y}$ 的且在 $[x_0,x_0+\Delta x]\subset\Omega$ 上具有连续导数的算子，则关于 $P'(x)$ 的积分存在，而且有如下的关系式

$$\int_{x_0}^{x_0+\Delta x} P'(x)\,dx = P(x_0+\Delta x) - P(x_0), \tag{18}$$

这个公式是积分学基本定理，即 Newton-Leibnitz 公式的推广.

事实上，

$$\int_{x_0}^{x_0+\Delta x} P'(x)\,dx = \lim_{\lambda\to 0}\sum_{k=0}^{n-1} P'(x_0+\tau_k\Delta x)\Delta x(t_{k+1}-t_k)$$

$$= \lim_{\lambda\to 0}\sum_{k=0}^{n-1} P'(\bar{x}_k)\Delta x_k$$

$(\bar{x}_k = x_0+\tau_k\Delta x;\ \Delta x_k = (t_{k+1}-t_k)\Delta x;\ k=0,1,\cdots,n-1)$. 另一方面，

$$P(x_0+\Delta x) - P(x_0) = \sum_{k=0}^{n-1}[P(x_0+t_{k+1}\Delta x) - P(x_0+t_k\Delta x)]$$

$$= \sum_{k=0}^{n-1}[P(x_{k+1}) - P(x_k)].$$

利用有限增量公式，得到

$$\left\|\sum_{k=0}^{n-1}[P(x_{k+1}) - P(x_k) - P'(\bar{x}_k)\Delta x_k]\right\|$$

$$\leqslant \|\Delta x\|\sum_{k=0}^{n-1}(t_{k+1}-t_k)\sup_{0<\theta<1}\|P'(x_k+\theta\Delta x_k) - P'(\bar{x}_k)\|,$$

由此，注意到 P' 的连续性，因而知 P' 在$[x_0, x_0+\Delta x]$上一致连续. 于是，我们得到了所要求的结果.

注. 由积分的性质 III 推知，如果

$$\|R(x_0+\tau\Delta x)\| \leqslant \varphi(t_0+\tau\Delta t) \qquad (\tau\in[0,1]) \tag{19}$$

且

$$\|\Delta x\| \leqslant \Delta t, \tag{20}$$

则

$$\left\|\int_{x_0}^{x_0+\Delta x} R(x)\,dx\right\| \leqslant \int_{t_0}^{t_0+\Delta t} \varphi(t)\,dt. \tag{21}$$

事实上,

$$\left\|\int_{x_0}^{x_0+\Delta x} R(x)\,dx\right\| = \left\|\int_0^1 R(x_0+\tau\Delta x)(\Delta x)\,d\tau\right\|$$

$$\leqslant \|\Delta x\| \int_0^1 \|R(x_0+\tau\Delta x)\|\,d\tau$$

$$\leqslant \Delta t \int_0^1 \varphi(t_0+\tau\Delta t)\,d\tau = \int_{t_0}^{t_0+\Delta t} \varphi(t)\,dt.$$

特别,由(21)推知,如果不等式

$$\|R(x)\| \leqslant \varphi(t) \tag{22}$$

对满足关系

$$\|x-x_0\| \leqslant t-t_0 \tag{23}$$

的所有 x 和 t 均成立,则

$$\left\|\int_{x_0}^{x_1} R(x)\,dx\right\| \leqslant \int_{t_0}^{t_1} \varphi(t)\,dt, \tag{24}$$

其中 x_1 是满足条件 $\|x_1-x_0\| \leqslant t_1-t_0$ 的任意元素.

事实上,令 $\Delta x = x_1 - x_0$, $\Delta t = t_1 - t_0$, $x = x_0 + \tau\Delta x$, $t = t_0 + \tau\Delta t$,即得到 $\|\Delta x\| \leqslant \Delta t$,而且

$$\|x-x_0\| = \tau\|\Delta x\| \leqslant \tau\Delta t = t-t. \qquad (\tau\in[0,1]),$$

也就是说,(23)式成立. 由(22)

$$\|R(x_0+\tau\Delta x)\| \leqslant \varphi(t_0+\tau\Delta t) \qquad (\tau\in[0,1]).$$

再利用(21)即得到(24).

§2. 二阶导数和双线性算子

2.1. 设将 Banach 空间 $\mathbf{X}$ 的开集合 Ω 映到 Banach 空间 $\mathbf{Y}$ 的算子 P 可导,即在 Ω 上 P' 存在. 正如上面指出的那样,P' 可看成是由 Ω 到算子空间 $B(\mathbf{X},\mathbf{Y})$ 的算子. 因此,言及 P' 在点 $x_0\in\Omega$ 处的导数是有意义的. 如果这个导数存在的话,则称为算子 P 的

二阶导数，并记作 $P''(x_0)$。如果 P' 可微，则称 P 二次可微。

如果在 Ω 的每个点上二阶导数存在，因而算子 P'' 有定义，它的导数叫做 P 的三阶导数。一般地，P 在 x_0 处的 n 阶导数 $P^{(n)}(x_0)$ 是定义为 $P^{(n-1)}$ 的导数。易见

$$P^{(n)}(x_0)\in B(\overbrace{\mathbf{X}, B(\mathbf{X}, \cdots, B(\mathbf{X}}^{n}, \mathbf{Y})\cdots)).$$

2.2. 空间 $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ 的元素直接的解释是不大直观的。为了能完全弄明白这一记号的实质，需要引进一些新的概念。

设 $\mathbf{X}$ 和 $\mathbf{Y}$ 是两个 Banach 空间而且每一对有序元素 x，$x'\in\mathbf{X}$ 和元素

$$y=B(x, x')\in\mathbf{Y}$$

相联系。二元算子 B 称为双线性算子，如果下述两个条件被满足：

I. 对任何 $x_1, x_2, x_1', x_2'\in\mathbf{X}$ 和数 α, β

$$B(\alpha x_1+\beta x_2, x')=\alpha B(x_1, x')+\beta B(x_2, x'),$$

$$(x, x'\in\mathbf{X}). \quad (1)$$

$$B(x, \alpha x_1'+\beta x_2')=\alpha B(x, x_1')+\beta B(x, x_2')$$

II. 存在这样一个实数 $M>0$，使得对任何 $x, x'\in\mathbf{X}$

$$\|B(x, x')\|\leqslant M\|x\|\|x'\|. \quad (2)$$

正如对线性算子的定义一样，称不等式(2)中 M 的最小值为算子 B 的范数并写作 $\|B\|$。易见

$$\|B\|=\sup_{\|x\|,\|x'\|\leqslant 1}|B(x, x')|. \quad (3)$$

不难验证，所有具有范数(3)的双线性算子的集合以自然的方式线性化后构成一个 Banach 空间，此空间我们记为 $B(\mathbf{X}^2, \mathbf{Y})$。

我们指出(不再详细地来建立了)，用类似的办法可以引进 n-线性算子的算子空间 $B(\mathbf{X}^n, \mathbf{Y})$*)。

---

*) 其细节见 Гаэурин[2] 和 Hille-Phillips 的著作。

现在考察双线性算子的一些例子.

在最简单的情况 $\mathbf{X}=\mathbf{Y}=\mathbf{R}^1$中, 双线性算子的一般形式是

$$B(x, x')=\alpha x x'$$

(其中 α 是实数).

现在设$\mathbf{X}$和$\mathbf{Y}$ 分别是m和n维空间. 以 $x_i(i=1, 2, \cdots, m)$表示空间$\mathbf{X}$中除了第 i 个坐标等于 1 外而其他坐标全等于零的元素. 元素 $y_j\in\mathbf{Y}(j=1, 2, \cdots, n)$ 也类似地定义. 我们来考察算子 $B\in B(\mathbf{X}^2, \mathbf{Y})$. 设

$$B(x_i, x_j)=(a_{ij}^{(1)}, a_{ij}^{(2)}, \cdots, a_{ij}^{(n)}) \qquad (i, j=1, 2, \cdots, m),$$

由(1), 对任意元素 $x, x'\in\mathbf{X}(x=(\xi_1,\xi_2,\cdots,\xi_m), x'=(\xi'_1, \xi'_2, \cdots, \xi'_m))$ 将有如下关系

$$y=B(x, x')=B(\sum_{i=1}^{m}\xi_i x_i, \sum^{m}\xi'_j x_j) = \sum_{i, j=1}^{m}\xi_i\xi'_j B(x_i, x_j).$$

因此, 如果 $y=(\eta_1, \eta_2, \cdots, \eta_n)$, 则

$$\eta_k=\sum_{i, j=1}^{m} a_{ij}^{(k)}\xi_i\xi'_j \qquad (k=1, 2, \cdots, n), \tag{4}$$

也就是说元素 y 的坐标是元素 x 和 x' 的坐标的双线性型.

显然, 不论具有三个指标的矩阵

$$(a_{ij}^{(k)}) \qquad (i, j=1, 2, \cdots, m; k=1, 2, \cdots, n) \tag{5}$$

怎样取, 按(4)式所定义的算子 B 总是双线性的.

因此, 在我们所考察的情形下, 每个线性算子确定了具有三个指标的矩阵(5)并被(5)所确定.

至于算子B的范数, 自然, 它依赖于空间 $\mathbf{X}$ 和 $\mathbf{Y}$ 的范数. 例如, 若 $\mathbf{X}=l_m^2, \mathbf{Y}=l_n^2$, 则由(4)再借助 Cauchy-Буняковский 不等式, 得到,

$$|\eta_k| = \left|\sum_{i,j=1}^{m} a_{ij}^{(k)}\xi_i\xi_j'\right|$$

$$\leqslant\left[\sum_{i=1}^{m}|\xi_i|^2\right]^{\frac{1}{2}}\left[\sum_{i=1}^{m}\left|\sum_{j=1}^{m}a_{ij}^{(k)}\xi_j'\right|^2\right]^{\frac{1}{2}}$$

$$\leqslant \Lambda_k\|x\|\|x'\|,$$

其中 Λ_k 表示矩阵 $A_kA_k^*$ 的最大特征值（参见V.2.8），A_k^* 表示 A_k 的转置矩阵，而

$$A_k = \begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1m}^{(k)} \\ a_{21}^{(k)} & a_{22}^{(k)} & \cdots & a_{2m}^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}^{(k)} & a_{m2}^{(k)} & \cdots & a_{mm}^{(k)} \end{pmatrix} \qquad (k=1, 2, \cdots, n).$$

因此

$$\|y\| = \left[\sum_{k=1}^{n}|\eta_k|^2\right]^{1/2} \leqslant \left[\sum_{k=1}^{n}\Lambda_k^2\right]^{1/2}\|x\|\|x'\|.$$

由此即给出下述估计

$$\|B\|\leqslant\left[\sum_{k=1}^{n}\Lambda_k^2\right]^{1/2}. \tag{6}$$

因为

$$\Lambda_k^2\leqslant\sum_{i,j=1}^{m}|a_{ij}^{(k)}|^2 \qquad (k=1, 2, \cdots, n),$$

所以由(6)可得到更粗的，然而是更简单的估计

$$\|B\|\leqslant\left[\sum_{k=1}^{n}\sum_{i=1}^{m}\sum_{j=1}^{m}|a_{ij}^{(k)}|^2\right]^{1/2}.$$

若令 $\mathbf{X}=l_m^\infty$，$\mathbf{Y}=l_n^\infty$，则经由极显然的论证可得对 $\|B\|$ 的下述估计

$$\| B \| \leqslant \max_k \sum_{i=1}^m \sum_{j=1}^m | a_{ij}^{(k)} |.$$

最有意义的是在无限维空间中，特别和主要地是在函数空间中双线性算子的例子.

算子 B:

$$y = B(x, x'), \quad y(s) = \int_0^1 \int_0^1 K(s, t, u) x(t) x'(t) dt du \tag{7}$$

根据对核函数 K 的要求的不同而可以在种种函数空间(定义在[0, 1]上的函数组成的函数空间）中是双线性的. 于是，若假定核函数连续，则不论 $\mathbf{X}$ 和 $\mathbf{Y}$ 取空间 C，$\mathbf{L}^p(1 \leqslant p \leqslant \infty)$ 中的哪个空间，B 的双线性性均将成立.

对应用最有价值的乃是形如

$$y = B(x, x') \tag{8}$$

$$y(s) = \int_0^1 K(s, t) x(t) x'(t) dt \tag{9}$$

的算子. 先设 $\mathbf{X} = \mathbf{Y} = \mathbf{L}^p(0, 1)(p \geqslant 2)$. 如果 $M < \infty$，其中

$$M^p = \begin{cases} \displaystyle\int_0^1 \left[\int_0^1 | K(s, t) |^{\frac{p}{p-2}} dt \right]^{p-2} ds & (p > 2), \\ \displaystyle\int_0^1 \sup_s | K(s, t) | dt & (p = 2), \end{cases}$$

则 $B \in B(\mathbf{X}^2, \mathbf{Y})$.

事实上，(9)式中积分号下的乘积对几乎所有的 $s \in [0, 1]$ 是可积的. 为了对 $p > 2$ 的情形验证这一事实，只须对积分(9)应用指数为 $\dfrac{p}{p-2}$，p，p 的广义 Hölder 不等式(Ⅳ.2.4)，依据这个不等式，我们得到

$$|y(s)|^p = \left|\int_0^1 K(s,t)x(t)x'(t)dt\right|^p$$

$$\leqslant \left[\int_0^1 |K(s,t)|^{\frac{p}{p-2}}dt\right]^{p-2}\int_0^1 |x(t)|^p dt \int_0^1 |x'(t)|^p dt.$$

因而

$$\int_0^1 |y(s)|^p ds \leqslant M^p \|x\|^p \|x'\|^p < \infty. \tag{10}$$

于是，$y \in \mathbf{L}^p(0,1)$. 同时，这还证明了：对于算子 B，不等式(2)是满足的. 由于可加性(关系式(1))，则显然可得到算子 B 的双线性性. 同时，由(10)式还推出范数的如下估计

$$\|B\| \leqslant M.$$

这就对 $p>2$ 的情形建立了我们的结论. 至于 $p=2$ 的情形也是不难得到的.

现在设 $\mathbf{X}=\mathbf{Y}=\mathbf{C}[0,1]$. 这时，算子(8)的双线性性乃是由核函数 $K(s,t)$ 关于 s 连续并且

$$M = \max_s \int_0^1 |K(s,t)|dt < \infty$$

而得以保证.

事实上，(9) 式中被积函数的可积性是不成问题的. y 的连续性可从能对积分(9) 积分号下取极限而推知. 等式(1) 是显然的. 同样容易验证不等式(2)成立，因为

$$y(s) \leqslant \max_t |x(t)| \max_t |x'(t)| \int_0^1 |K(s,t)|dt \qquad (0 \leqslant s \leqslant 1),$$

因此

$$\|y\| \leqslant M\|x\|\|x'\|.$$

由此得到估计

$$\|B\| \leqslant M = \sup_s \int_0^1 |K(s, t)| dt,$$

而实际上是等号成立:

$$\|B\| = \sup_s \int_0^1 |K(s, t)| dt$$

(这可以象在 V.2.4 中对线性算子的相应的等式那样来确立).

2.3. 上面所引入的空间 $B(\mathbf{X}^2, \mathbf{Y})$ 与我们在二阶导数的讨论中所遇到的空间 $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ 有紧密的联系. 亦即, 我们来证明这两个空间是线性等距的.

设 $W \in B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$. 取任意元素 $x' \in \mathbf{X}$ 并令 $U_{x'} = W(x')$. 显然, $U_{x'} \in B(\mathbf{X}, \mathbf{Y})$, 因此 $y = B(x, x') = U_{x'}(x)$ 是空间 $\mathbf{Y}$ 中的元素. 这样定义的二元算子 B 显然满足 2.2 中的条件 I, 又由于

$$\|y\| \leqslant \|U_{x'}\| \|x\| \leqslant \|W\| \|x\| \|x'\|, \tag{11}$$

所以条件 II 同样是满足的.

于是, $B \in B(\mathbf{X}^2, \mathbf{Y})$, 并且由(11)推出不等式

$$\|B\| \leqslant \|W\|. \tag{12}$$

可见, 每个算子 $W \in B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ 均可用这种方法与双线性算子 $B \in B(\mathbf{X}^2, \mathbf{Y})$ 联系起来. 我们来证明, 在这样的对应下, 每个双线性算子 $B \in B(\mathbf{X}^2, \mathbf{Y})$ 都是某个算子 $W \in B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ 的象.

为此, 只须指出: 如果 B 是给定的, 则当固定 $x' \in \mathbf{X}$ 时, 算子

$$W(x') = B(\cdot, x') \in B(\mathbf{X}, \mathbf{Y})^{*)}.$$

在这种情况下

$$\|W(x')\| = \sup_{\|x\| \leqslant 1} \|W(x')(x)\| = \sup_{\|x\| \leqslant 1} \|B(x, x')\| \leqslant \|B\| \|x'\|,$$

---

*) $B(\cdot, x')$ 表示使得 $U(x) = B(x, x')$ 这样的算子 U. 类似的表示法在以后还要用到.

由此推知 $W\in B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$，并且有

$$\|W\|\leqslant\|B\|.$$

将这个不等式和(12)相比较，则得到等式

$$\|W\|=\|B\|. \tag{13}$$

因为空间 $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$和$B(\mathbf{X}^2, \mathbf{Y})$的元素间的对应是可加的和齐次的，所以从(13)式推知这两个空间是一对一的，这样，就实现了它们之间的线性等距关系.

空间 $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$和 $B(\mathbf{X}^2, \mathbf{Y})$的线性等距为这两个空间对应元素的等同给出了依据，我们在下面就是这样看待的.

我们不加证明地指出(其实读者不难自己去验证)：在一般情况下，空间 $B(\mathbf{X}, \cdots, B(\mathbf{X}, \mathbf{Y}), \cdots)$ 和n-线性算子空间 $B(\mathbf{X}^n, \mathbf{Y})$ 是线性等距的(参见 Гавурин[2]).

2.4. 正如在 2.1 中所指出的，从 $\mathbf{X}$ 到 $\mathbf{Y}$ 的算子P的二阶导数乃是空间 $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$中的元素，由此，与上面所得到的结果相对照，即可将 $P''(x_0)$看成双线性算子. 在将 $P''(x_0)$ 看成双线性算子的情况下，按照 2.3 所述，将有

$$\lim_{t\to 0}\frac{P'(x_0+tx')-P'(x_0)}{t}=P''(x_0)(\cdot, x'). \tag{14}$$

因而，对任意的 $x\in\mathbf{X}$，

$$P''(x_0)(x, x')=\lim_{t\to 0}\frac{P'(x_0+tx')x-P'(x_0)x}{t}. \tag{15}$$

应该注意，等式(14)和(15)并不等价. 更确切地说，可能存在双线性算子 $B\in B(\mathbf{X}^2, \mathbf{Y})$，使得对任何 $x, x'\in\mathbf{X}$ 有

$$\lim_{t\to 0}\frac{P'(x_0+tx')x-P'(x_0)x}{t}=B(x, x') \tag{16}$$

而与此同时 $P''(x_0)$却不存在. 不难看出，为了使得 $P''(x_0)$ 存在的必要充分条件是：对每个 $x'\in\mathbf{X}$，极限关系(16)对于$\|x\|=1$的所有 $x\in\mathbf{X}$ 一致地成立. 所指出的这一情况使得可以对二阶导数

的概念作若干扩充. 也就是说，如果对任意的 $x, x'\in\mathbf{X}$, (16)均成立，则称双线性算子 B 是算子 P 在 x_0 点的弱二阶导数. 显然，对于弱导数仍然可以沿用以前的表示，这是因为若二阶导数在通常意义下存在的话，它与弱二阶导数是一致的.

注. 不难看到，如果 P 在点 x_0 是二次可微的，则这表示(16)式对于 $\|x\|=1$ 和 $\|x'\|=1$ 的所有 $x, x'\in\mathbf{X}$ 是一致地成立的.

我们现在来弄清楚当 $\mathbf{X}$ 和 $\mathbf{Y}$ 是有限维空间时，二阶导数具何种形式.

设 $P''(x_0)$ 存在而且它作为双线性算子是由三个指标的矩阵

$$(a_{ij}^{(k)}) \qquad (i, j=1, 2, \cdots, m; k=1, 2, \cdots, n)$$

所确定的. 这时，比较关系式 (15)左、右两端的对应坐标并沿用1.6 中的记法，得到

$$\sum_{i,j=1}^{m} a_{ij}^{(k)}\xi_i\xi'_j=\lim_{t\to 0}\frac{\sum\limits_{i=1}^{m}\frac{\partial\varphi_k}{\partial\xi_i}(\xi_j^{(0)}+t\xi'_j)\xi_i-\sum\limits_{i=1}^{m}\frac{\partial\varphi_k}{\partial\xi_i}(\xi_j^{(0)})\xi_i}{t}$$

$$(k=1, 2, \cdots, n)^{*)}.$$

作为 x 和 x' 我们这样来选取，使得除 x 的第 i 个坐标与 x' 的第 j 个坐标都等于1以外，其余所有的坐标都等于零，这样，我们就得到函数 φ_k 具有二阶导数而且

$$a_{ij}^{(k)}=\frac{\partial^2\varphi_k(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_m^{(0)})}{\partial\xi_i\partial\xi_j}$$

$$(i, j=1, 2, \cdots, m;\ k=1, 2, \cdots, n).$$

2.5. 我们在这一段里来建立一个公式. 这个公式乃是对于通常实函数的 Taylor 公式的推广.

象以前一样，设 P 是从空间 $\mathbf{X}$ 的集合 Ω 作用到空间 $\mathbf{Y}$ 的算

---

*) 这里采用了简化记号

$$\varphi(\xi_j)=\varphi(\xi_1, \xi_2, \cdots, \xi_m).$$

子，而且它在区间$[x_0, \bar{x}]\subset\Omega$上具有连续的二阶导数．这时成立如下的公式：

$$P(\bar{x})-P(x_0)-P'(x_0)(\bar{x}-x_0)=\int_{x_0}^{\bar{x}} P''(x)(\bar{x}-x,\cdot)dx. \quad (17)$$

为证明等式(17)，我们来考察算子Q：

$$Q(x)=P(x)+P'(x)(\bar{x}-x)=P(x)+Q_0(x) \qquad (x\in\Omega)$$

并检验对任何$\tilde{x}\in[x_0, \bar{x}]$

$$Q'(\tilde{x})(x)=P''(\tilde{x})(\bar{x}-\tilde{x}, x) \qquad (x\in\mathbf{X}). \qquad (18)$$

如果这一点得以确立，则公式(17)显然可从 Newton-Leibnitz 公式(见1.7 的(18)式)得到．

只须考察算子Q_0．我们有

$$\frac{Q_0(\tilde{x}+tx)-Q_0(\tilde{x})}{t}=\frac{P'(\tilde{x}+tx)(\bar{x}-\tilde{x}-tx)-P'(\tilde{x})(\bar{x}-\tilde{x})}{t}$$

$$=\frac{P'(\tilde{x}+tx)(\bar{x}-\tilde{x})-P'(\tilde{x})(\bar{x}-\tilde{x})}{t}-P'(\tilde{x}+tx)x.$$

当$t\to 0$时，第一项具有极限$P''(\tilde{x})(\bar{x}-\tilde{x}, x)$，第二项由于算子$P'$的连续性而趋于$P'(\tilde{x})(x)$．因此

$$Q_0'(\tilde{x})(x)=P''(\tilde{x})(\bar{x}-\tilde{x}, x)-P'(\tilde{x})(x),$$

这就对于$Q'(\tilde{x})(x)$给出了关系式(18)．

注．由以上结论易知，若$P''(x)$按弱二阶导数来理解，公式(17)亦正确．

§3. 例　子

在上一节中，我们业已遇到了在最简单的情况下，即在有限维空间情况下求非线性算子的导数的例子．在这一节里，我们将要考察更复杂和更富有内容的一些例子，这些例子本身将在下一节并且主要是在下一章中得到应用．

3.1. 我们来考察积分算子P：

$$y = P(x), \qquad y(s) = \int_0^1 K(s, t, x(t))dt$$

定理 1. 设函数 $K(s, t, u)$ 对于 $0 \leqslant s, \ t \leqslant 1, \ -\infty < u < +\infty$ 连续，而且关于 u 二次连续可微，而且对 s, t, u 的这些值

$$|K''_{u^2}(s, t, u)| \leqslant M|u|^{p-2} + N. \tag{2}$$

那么，若 $P \geqslant 2$，则算子(1)映空间 $\mathbf{L}^p(0, 1)$ 到空间 $\mathbf{L}^q(1 \leqslant q < \infty)$，而且在每个点 $x_0 \in \mathbf{L}^p(0, 1)$ 可微并具有二阶导数. 在这种情况下，$P'(x_0) = U$，$P''(x_0) = B$，其中 U 和 B 由下述公式所确定：

$$z = U(x), \qquad z(s) = \int_0^1 K'_u(s, t, x_0(t))x(t)dt,$$

$$v = B(x, x'), \quad v(s) = \int_0^1 K''_{u^2}(s, t, x_0(t))x(t)x'(t)dt.$$

证. 首先，可以确信(1)是映 $\mathbf{L}^p(0, 1)$ 到 $\mathbf{L}^\infty(0, 1)$（因而亦是到空间 $\mathbf{L}^q(0, 1)(q < \infty)$）的算子. 为此目的，只须指出：由于 $K(s, t, u)$ 是连续的，所以复合函数 $K(s, t, x(t))$ 是可测的.

其次，依 Taylor 公式，有

$$K(s, t, x(t)) = K(s, t, 0) + K'_u(s, t, 0)x(t)$$
$$+ \frac{1}{2}K''_{u^2}(s, t, \theta x(t))x^2(t) \qquad (0 < \theta < 1).$$

我们现在来对等式右端每一项在[0, 1]上的积分作估计：

$$\left|\int_0^1 K(s, t, 0)dt\right| \leqslant \max_{s,t}|K(s, t, 0)| = A_1,$$

$$\left|\int_0^1 K'_u(s, t, 0)x(t)dt\right| \leqslant \max_{s,t}|K'_u(s, t, 0)|\int_0^1 |x(t)dt$$
$$\leqslant A_2\|x\|_{\mathbf{L}^p}.$$

利用(2)，可对第三项的积分作如下估计：

$$\left|\frac{1}{2}\int_0^1 K''_{u^2}(s, t, \theta x(t))x^2(t)dt\right|$$
$$\leqslant \frac{1}{2}\int_0^1 [M|x(t)|^{p-2} + N]|x(t)|^2 dt \leqslant A_3\|x\|^p_{\mathbf{L}^p}.$$

由此可知 $y=P(x)$ 有意义并且有如下估计

$$|y(s)|\leqslant A_1+A_2\|x\|_{\mathbf{L}^p}+A_3\|x\|_{\mathbf{L}^p}^p \qquad (s\in[0,1]),$$

因此 $y\in\mathbf{L}^\infty(0,1)$.

现在来证明 P 是可微的，并且 $P'(x_0)=U$. 设 $z=U(x)$ 而

$$z_\tau=\frac{P(x_0+\tau x)-P(x_0)}{\tau}.$$

于是就有

$$\begin{aligned}&|z_\tau(s)-z(s)|\\&=\left|\int_0^1\left[\frac{K(s,t,x_0(t)+\tau x(t))-K(s,t,x_0(t))}{\tau}\right.\right.\\&\left.\left.-K'_u(s,t,x_0(t))x(t)\right]dt\right|.\end{aligned}\qquad(3)$$

根据 Taylor 公式

$$\begin{aligned}&K(s,t,x_0(t)+\tau x(t))=K(s,t,x_0(t))\\&+\tau K'_u(s,t,x_0(t))x(t)+\frac{\tau^2}{2}K''_{u^2}(s,t,x_0(t)+\theta\tau x(t))x^2(t)\\&\qquad(0<\theta<1),\end{aligned}$$

可将(3)改写为

$$|z_\tau(s)-z(s)|=\frac{|\tau|}{2}\left|\int_0^1 K''_{u^2}(s,t,x_0(t)+\theta\tau x(t))dt\right|.$$

利用与上面完全类似的讨论，可得到估计

$$|z_\tau(s)-z(s)|\leqslant A_4|\tau| \qquad (0\leqslant s\leqslant 1).$$

其中常数 A_4 仅仅依赖于数 $\|x\|$ 而不依赖于 x 本身. 因此，当 $\tau\to 0$ 时，对于 $\|x\|=1$ 的所有 x 一致地有 $z_\tau\to z$.

现在来求二阶导数. 记

$$v=B(x,x'); \qquad v_\tau=\frac{P'(x_0+\tau x')(x)-P'(x_0)(x)}{\tau}.$$

正如在 2.4 中所指出的那样，只要证明当 $\tau\to 0$ 时，对于 $\|x\|=1$ 的所有的 $x\in\mathbf{L}^p(0,1)$ 一致地有

$$v_\tau \longrightarrow v$$

就够了.

先应用有限增量公式，再利用 Hölder 不等式，得到

$$
\begin{aligned}
&|v_\tau(s)-v(s)| \\
&=\left|\int_0^1\left[\frac{K'_u(s,t,x_0(t)+\tau x'(t))-K'_u(s,t,x_0)}{\tau}\right.\right. \\
&\qquad \left.\left. -K''_{u^2}(s,t,x_0(t))x'(t)\right]dt\right| \\
&=\left|\int_0^1[K''_{u^2}(s,t,x_0(t)+\tau\theta x'(t))\right. \\
&\qquad\qquad \left. -K''_{u^2}(s,t,x_0(t))]x(t)x'(t)\,dt\right| \\
&\leqslant\|x\|\left\{\int_0^1|[K''_{u^2}(s,t,x_0(t)+\theta\tau x'(t))\right. \\
&\qquad\qquad \left.-K''_{u^2}(s,t,x_0(t))]x'(t)|^{\frac{p}{p-1}}dt\right\}^{\frac{p-1}{p}}.
\end{aligned}
\tag{4}
$$

如果 $x_0(t)$ 和 $x'(t)$ 是有限的，则最后一个积分中被积函数当 $\tau\to 0$ 时趋于零. 另外，由(2)式

$$
\begin{aligned}
&|[K''_{u^2}(s,t,x_0(t)+\tau\theta x'(t))-K''_{u^2}(s,t,x_0(t))]x'(t)|^{\frac{p}{p-1}} \\
&\leqslant\{[M|x_0(t)+\theta\tau x'(t)|^{p-2}+M|x_0(t)|^{p-2}+2N]|x'(t)|\}^{\frac{p}{p-1}} \\
&\qquad\leqslant A_5|x_0(t)|^{\frac{(p-2)p}{p-1}}|x'(t)|^{\frac{p}{p-1}}+A_6|x'(t)|^{p\,*)}.
\end{aligned}
$$

由于乘积 $|x_0(t)|^{p-\frac{p}{p-1}}|x'(t)|^{\frac{p}{p-1}}$ 和 $|x'(t)|^p$ 是可积的，故可在(4)中积分号下取极限，这就给出

$$\lim_{\tau\to 0}v_\tau(s)=v(s)\qquad(0\leqslant s\leqslant 1);$$

由此显然可见，对于 $\|x\|=1$ 的所有 x 来说 v_τ 是一致地趋于 v 的.

由以上的讨论同样可推知

$$|v_\tau(s)|\leqslant\int_0^1|K''_{u^2}(s,t,x_0(t)+\theta\tau x'(t))x(t)x'(t)|dt\leqslant A_7,$$

---

*) 应利用当 $\alpha>0$ 时的不等式 $|a+b|^\alpha\leqslant 2^\alpha(|a|^\alpha+|b|^\alpha)$.

由此可在积分

$$\|v_\tau - v\|_{L^q} = \left\{\int_0^1 |v_\tau(s) - v(s)|^q ds\right\}^{1/q}$$

中于积分号下取极限，也就是说$\|v_\tau - v\|$关于$\|x\|=1$的所有x是一致地趋于零的.

定理获证.

注. 如果将条件(2)变为条件

$$|K''_{u^2}(s,t,u)| \leqslant M|u|^\lambda + N \qquad (\lambda < p-2) \tag{5}$$

并且如果$p>2$，则可以证明：算子P(视作从$\mathbf{L}^p$到$\mathbf{L}^\infty$的算子)在每个点$x_0 \in \mathbf{L}^p$是二次可微的，这里的x_0是有界函数.

3.2. 算子(1)还可以看作是从$\mathbf{C}[0,1]$到$\mathbf{C}[0,1]$的算子.

设x_0是空间$\mathbf{C}[0,1]$中的一固定元素. 以Ω表示空间$\mathbf{C}[0,1]$中以x_0为中心以r为半径的球，以G表示由不等式

$$0 \leqslant s,\quad t \leqslant 1;\quad |u - x_0(t)| \leqslant r$$

所定义的三维空间中点(s,t,u)的集合.

定理 2. 设函数$K(s,t,u)$在G上给定并且连同导数$K'_u(s,t,u)$和$K''_{u^2}(s,t,u)$在这个集合上是连续的. 那么，由公式(1)所定义的P是从集合Ω到$\mathbf{C}[0,1]$的一个算子，它在每个内点$\bar{x} \in \Omega$是二次可微的，并且$P'(\bar{x}) = U$, $P''(\bar{x}) = B$，其中U和B由下述公式确定：

$$z = U(x), \qquad z(s) = \int_0^1 K'_u(s,t,\bar{x}(t))x(t)dt,$$

$$v = B(x, x'), \qquad v(s) = \int_0^1 K''_{u^2}(s,t,\bar{x}(t))x(t)x'(t)dt.$$

证. 从Ω中任取x. 由公式(1)所定义的函数$y = P(x)$的连续性可直接从依赖于参数的积分论的基本事实推知，这就表明$y \in \mathbf{C}[0,1]$.

现在设$\bar{x} \in \Omega$, $x \in \mathbf{C}[0,1]$. 考察元素

$$z_\tau = \frac{P(\bar{x}+\tau x)-P(\bar{x})}{\tau} - U(x).$$

于是就有

$$z_\tau(s)=\int_0^1\left[\frac{K(s,t,\bar{x}(t)+\tau x(t))-K(s,t,\bar{x}(t))}{\tau}\right.$$

$$\left. -K'_u(s,t,\bar{x}(t))x(t)\right]dt.$$

当 $\tau\to 0$ 时，被积表达式趋于零，而且对于 $\|x\|=1$ 的所有 x 和 $s\in[0,1]$ 是一致有界的（这一点可由有限增量公式确认）. 因此，z_τ 对 $\|x\|=1$ 的所有的 $x\in \mathbf{C}[0,1]$ 一致地趋于零. 这就证得了 P 在点 $\bar{x}$ 的可微性.

此外，

$$P'(\bar{x})=U.$$

用类似的办法可证 $P''(\bar{x})=B$ 而且 P 是二次可微的.

3.3. 我们来考察与一阶常微分方程组

$$\frac{dy_i}{ds}=f_i(s,y_k(s)) \qquad (i,k=1,2,\cdots,n)$$

相联系的非线性算子. 为方便计可将此方程组写成积分方程组的形式

$$y_i(s)=y_i^{(0)}+\int_0^s f_i(t,y_k(t))\,dt \qquad (i,k=1,2,\cdots,n)$$

或者可更简略地写为向量形式

$$y(s)=y^{(0)}+\int_0^s f(t,y(t))\,dt,$$

其中 $y(s)$, $y^{(0)}$, $f(t,u)$ 和 u 是分别由

$$(y_1(s),y_2(s),\cdots,y_n(s)),\quad (y_1^{(0)},y_2^{(0)},\cdots,y_n^{(0)}),$$

$$(f_1(t,u_k),\ f_2(t,u_k),\cdots,f_n(t,u_k)),\ (u_1,u_2,\cdots,u_n)$$

所组成的向量. 我们来考察空间 $\mathbf{C}_n^\alpha$ $(\alpha>0)$，这个空间的元素乃是 $[0,\infty)$ 上绝对连续的向量函数，而且

$$\sup_t |y(t)e^{\alpha t}| < \infty, \quad \sup_t |y'(t)e^{\alpha t}| < \infty \quad (t \geqslant 0)^{*)}.$$

空间 $\mathbf{C}_n^\alpha$ 中的范数借助于等式

$$\|y\| = \sup_t |y(t)e^{\alpha t}| + \sup_t |y'(t)e^{\alpha t}| \qquad (y \in \mathbf{C}_n^\alpha)$$

来引进.

除了 $\mathbf{C}_n^\alpha$以外，我们还来考察空间 $\mathbf{D}_n^\alpha$，这个空间是由满足条件

$$\sup_t |z(t)| < \infty, \quad \sup_t |z'(t)e^{\alpha t}| < \infty \qquad (z \in \mathbf{D}_n^\alpha)$$

的绝对连续的向量函数组成的. $\mathbf{D}_n^\alpha$ 中的范数定义为:

$$\|z\| = \sup_t |z(t)| + \sup_t |z'(t)e^{\alpha t}| \qquad (z \in \mathbf{D}_n^\alpha).$$

不难验证 $\mathbf{C}_n^\alpha$ 和 $\mathbf{D}_n^\alpha$ 都是 Banach 空间.

我们来考察算子 P:

$$z = P(y), \tag{6}$$

$$z(s) = y(s) - \int_0^s f(t, y(t))\,dt. \tag{7}$$

定理 3. 设 $f(t,u)$ 是定义在 $(n+1)$维空间的集合 G 上的向量函数，其中 G 是这样定义的: $(t,u) \in G$ 乃意味着 $t \geqslant 0$, $|u| \leqslant re^{-\alpha t}$. 同时

1) $f(t,0) = 0$;

2) 在 G 中存在有界导数 $f'_u(t,u)$**)，此导数对 $t \geqslant 0$关于 u 是一致连续的.

这时，如果以 Ω 表示空间 $\mathbf{C}_n^\alpha$ 中以零为中心以 r 为半径的球，则算子(6)将 Ω 变到 $\mathbf{D}_n^\alpha$ 且在 Ω 的每一点都是可微的，而且 $P'(y_0) = U$，其中 U 由下述公式确定:

---

*) 正象在第十六章中一样，向量 u 的长度用 $|u|$ 表示，而视为从 $\mathbf{l}_n^2$ 到 $\mathbf{l}_n^2$ 的算子的矩阵 A 的范数用 $|A|$ 表示.

**) 我们记得，向量函数 $f(t,\cdot)$ 的导数 $f'_u(t,u)$ 是导数的 n 阶方阵 $\left(\frac{\partial f_i}{\partial u_k}\right)$ (参见 1.6)

$$z=U(y),\quad z(s)=y(s)-\int_0^1 f'_u(t,y_0(t))y(t)\,dt. \qquad (8)$$

证. 首先来证明由等式(7)定义的函数属于 $\mathbf{D}_n^\alpha$(假定 $y\in\Omega$).事实上,由于 $f(t,0)=0$,所以

$$z(s)=y(s)-\int_0^s [f(t,y(t))-f(t,0)]\,dt.$$

这里的被积表达式可借助于有限增量公式用下述方法来估计:

$$|f(t,y(t))-f(t,0)|\leqslant \sup_{0<\theta<1}|f'_u(t,\theta y(t))|\,|y(t)|$$

$$\leqslant M\|y\|e^{-\alpha t}$$

$$(M=\sup|f'_u(t,u)|,\quad (t,u)\in G).$$

由此得到

$$\sup_s|z(s)|\leqslant\sup_s|y(s)|+M\|y\|\sup_s\int_0^s e^{-\alpha t}\,dt$$

$$\leqslant\|y\|+\frac{M}{\alpha}\|y\|<\infty.$$

而且

$$z'(s)=y'(s)-f(s,y(s)),$$

按类似的讨论可确认

$$\sup_s|z'(s)e^{\alpha s}|<\infty.$$

因此可知 z 有意义而且是空间 $\mathbf{D}_n^\alpha$ 的元素.

现在来验证由(8)式所定义的 U 是映 $\mathbf{C}_n^\alpha$ 到 $\mathbf{D}_n^\alpha$ 的连续线性算子.

重复以上的讨论,对 $z=U(y)$ 得到

$$\sup_s|z(s)|+\sup_s|z'(s)e^{\alpha s}|$$

$$=\sup_s\left|y(s)-\int_0^s f'_u(t,y_0(t))y(t)\,dt\right|$$

$$+\sup_s|[y'(s)-f'_u(s,y_0(s))y(s)]e^{\alpha s}|\leqslant$$

$$\leqslant \|y\| + \frac{M}{\alpha}\|y\| + (M+1)\|y\| = L\|y\|,$$

由此可知 $z = U(y) \in \mathbf{D}_n^\alpha$ 而且 $\|z\| \leqslant L\|y\|$.

如同在上一定理中所做的那样，我们现在来考察元素

$$v_\tau = \frac{P(y_0 + \tau y) - P(y_0)}{\tau} - U(y), \quad v_\tau(s) = \int_0^s J_\tau(t)\,dt$$

$$\left(J_\tau(t) = \frac{f(t, y_0(t) + \tau y(t)) - f(t, y_0(t))}{\tau} - f'_u(t, y_0(t))y(t)\right).$$

应用有限增量公式于被积表达式，得到

$$|J_\tau(t)| \leqslant \sup_{0<\theta<1} |f'_u(t, y_0(t) + \theta\tau y(t)) - f'_u(t, y_0(t))|\,|y(t)|.$$

因此，由于 $f'_u(t, u)$ 的一致连续性，可知对无论怎样的 $\varepsilon > 0$，当 τ 充分小时，对于 $\|y\| = 1$ 的所有的 y 一致地有

$$|J_\tau(t)| \leqslant \varepsilon e^{-\alpha t},$$

因而

$$\sup_s |v_\tau(s)| \leqslant \frac{\varepsilon}{\alpha}.$$

类似地，对于这些 τ 将有

$$\sup_s |v'_\tau(s) e^{\alpha s}| \leqslant \varepsilon,$$

由此得到

$$\|v_\tau\| \leqslant \varepsilon\left(\frac{1}{\alpha} + 1\right).$$

因此，当 $\tau \to 0$ 时对于 $\|y\| = 1$ 的所有的 $y \in \mathbf{C}_n^\alpha$ 来说 v_τ 一致地趋于

$\mathbf{X}\times\mathbf{Y}$ 依自然的方式赋予线性运算，并用等式

$$\|(x,y)\|=\|x\|+\|y\|$$

在其中引入范数，则 $\mathbf{X}\times\mathbf{Y}$ 也构成 Banach 空间. 这个 Banach 空间称为空间 $\mathbf{X}$ 和 $\mathbf{Y}$ 的直积空间(参见 IV.1.8).

将形如$(x,\mathbf{0_Y})$的元素偶与空间 $\mathbf{X}$ 的元素视为等同元素，则可将 $\mathbf{X}$ 看成直积空间 $\mathbf{X}\times\mathbf{Y}$ 的子空间，其中 $x\in\mathbf{X}$ 而$\mathbf{0_Y}$是空间 $\mathbf{Y}$ 中的零元素. 当然，对空间 $\mathbf{Y}$ 也可作类似的理解. 在这种理解下，每一个元素 $u=(x,y)\in\mathbf{X}\times\mathbf{Y}$ 均可唯一地表示成

$$u=(x,y)=(x,\mathbf{0_Y})+(\mathbf{0_X},y)=x+y.$$

每一个线性算子 $U\in B(\mathbf{X}\times\mathbf{Y},\mathbf{Z})$ 都诱导出一线性算子偶$(U_{\mathbf{X}},U_{\mathbf{Y}})$，其中 $U_{\mathbf{X}}\in B(\mathbf{X},\mathbf{Z})$，$U_{\mathbf{Y}}\in B(\mathbf{Y},\mathbf{Z})$而且

$$U_{\mathbf{X}}(x)=U((x,\mathbf{0_Y})),\quad U_{\mathbf{Y}}(y)=U((\mathbf{0_X},y)).$$

反之，如果存在算子偶$(U_{\mathbf{X}},U_{\mathbf{Y}})$，则算子 U:

$$U((x,y))=U_{\mathbf{X}}(x)+U_{\mathbf{Y}}(y)$$

将是从空间 $\mathbf{X}\times\mathbf{Y}$ 到 $\mathbf{Z}$ 的连续线性算子.

现在设 P 是映集合 $\Omega\subset\mathbf{X}\times\mathbf{Y}$ 到 Banach 空间 $\mathbf{Z}$ 的非线性算子. 固定 $y_0\in\mathbf{Y}$，我们来考察算子

$$P^{(y_0)}=P(\cdot,y_0),$$

这是将集合 $\Omega^{(y_0)}$映射到空间 $\mathbf{Z}$ 中的算子，其中 $\Omega^{(y_0)}$ 是元素 $x\in\mathbf{X}$ 的集合而$(x,y_0)\in\Omega$(即 $\Omega^{(y_0)}$ 是集合 Ω 的“截面”). 用类似的方法也可定义映 $\Omega^{(x_0)}\subset\mathbf{Y}$ 到 $\mathbf{Z}$ 的算子 $P^{(x_0)}(x_0\in\mathbf{X})$.

如果 $x_0\in\mathbf{X}$ 是集合 $\Omega^{(y_0)}$的内点，则可谈及关于导数$P^{(y_0)\prime}(x_0)$的问题. 这个导数称为算子 P 在点(x_0,y_0)处关于 x 的偏导数，并记作$P'_x(x_0,y_0)$. 类似地还可引入关于 y 的偏导数 $P'_y(x_0,y_0)$. 显然，

$$P'_x(x_0,y_0)\in B(\mathbf{X},\mathbf{Z}),\ P'_y(x_0,y_0)\in B(\mathbf{Y},\mathbf{Z}).$$

如果算子 P 在点$(x_0,y_0)\in\Omega$ 有导数 $P'(x_0,y_0)=U$，则 $P'_x(x_0,$

$y_0)=U_{\mathbf{X}}$ 而 $P'_y(x_0,y_0)=U_{\mathbf{Y}}$；这一情况表明所引入的偏导数记号是正确的.

算子 P 可看作抽象的"二个变元的函数"，将这种理解反映到记法上，即可用写法 $P(x,y)$ 去代替 $P((x,y))$. 关于多变元函数初等理论的某些基本原理不难推广到这样的"函数"上来. 特别是我们要指出其中的一个基本事实.

假定在点 $(x_0,y_0)\in\Omega$ 的某个邻域内存在导数 P'_x 和 P'_y，则如下不等式成立：

$$\begin{aligned}&\|P(x_0+\Delta x,y_0+\Delta y)-P(x_0,y_0)-P'_x(x_0,y_0)(\Delta x)\\&\qquad\qquad-P'_y(x_0,y_0)(\Delta y)\|\\&\leqslant\sup_{0<\theta,\theta_1<1}\|P'_x(x_0+\theta\Delta x,y_0+\theta_1\Delta y)-P'_x(x_0,y_0)\|\|\Delta x\|\\&+\sup_{0<\theta,\theta_1<1}\|P'_y(x_0+\theta\Delta x,y_0+\theta_1\Delta y)-P'_y(x_0,y_0)\|\|\Delta y\|.\quad(1)\end{aligned}$$

事实上，显然有

$$\begin{aligned}A=&\|P(x_0+\Delta x,y_0+\Delta y)-P(x_0,y_0)-P'_x(x_0,y_0)(\Delta x)\\&\qquad-P'_y(x_0,y_0)(\Delta y)\|\\\leqslant&\|P(x_0+\Delta x,y_0+\Delta y)-P(x_0,y_0+\Delta y)-P'_x(x_0,y_0)(\Delta x)\|\\&+\|P(x_0,y_0+\Delta y)-P(x_0,y_0)-P'_y(x_0,y_0)(\Delta y)\|.\end{aligned}$$

对不等式右端每一项应用有限增量公式(见§1,(8))，得到

$$\begin{aligned}A\leqslant&\sup_{0<\theta<1}\|P'_x(x_0+\theta\Delta x,y_0+\Delta y)-P'_x(x_0,y_0)\|\|\Delta x\|\\&+\sup_{0<\theta_1<1}\|P'_y(x_0,y_0+\theta_1\Delta y)-P'_y(x_0,y_0)\|\|\Delta y\|\\\leqslant&\sup_{0<\theta,\theta_1<1}\|P'_x(x_0+\theta\Delta x,y_0+\theta_1\Delta y)-P'_x(x_0,y_0)\|\|\Delta x\|\\&+\sup_{0<\theta,\theta_1<1}\|P'_y(x_0+\theta\Delta x,y_0+\theta_1\Delta y)-P'_y(x_0,y_0)\|\|\Delta y\|,\end{aligned}$$

这就是需要证明的.

4.2. 这一节的基本结果和方程

$$P(x,y)=0$$

的解的存在性问题有关，也就是说和由这个方程定义的隐函数的存在问题有关．正象在初等分析学中那样，如下的定理成立．

定理 1. 设在点$(x_0, y_0)\in \mathbf{X}\times\mathbf{Y}$的某邻域$\Omega$上给定了算子$P$，它映$\Omega$到空间$\mathbf{Z}$并在点$(x_0, y_0)$处连续．此外，如果下述条件满足：

1) $P(x_0, y_0)=\mathbf{0}$;

2) 偏导数P'_y在Ω中存在，在点(x_0, y_0)处连续；

3) 算子$P'_y(x_0, y_0)\in B(\mathbf{Y}, \mathbf{Z})$具有连续逆算子

$$\Gamma=[P'_y(x_0, y_0)]^{-1}\in B(\mathbf{Z}, \mathbf{Y}),$$

则存在从点x_0的某邻域$G\subset\mathbf{X}$到空间$\mathbf{Y}$的算子F，具有如下性质：

a) $P(x, F(x))=\mathbf{0} \qquad (x\in G)$;

b) $F(x_0)=y_0$;

c) F在点x_0处连续．

由性质 a)—c) 所确定的算子F在下述意义下是唯一的，即：如果F_1是具有这三个性质的另一算子，则存在这样的$\eta>0$使得当$\|x-x_0\|<\eta$时

$$F_1(x)=F(x).$$

证．不失一般性，可以认为$x_0=\mathbf{0}_\mathbf{X}$, $y_0=\mathbf{0}_\mathbf{Y}$．取$x\in\mathbf{X}$充分小使得$\mathbf{0}_\mathbf{Y}\in\Omega^{(x)}$，这里的$\Omega^{(x)}$和以前一样表示使$(x, y)\in\Omega$的$y\in\mathbf{Y}$的集合．我们来考察定义在$\Omega^{(x)}$上的算子$Q^{(x)}$：

$$Q^{(x)}(y)=y-\Gamma P(x, y).$$

现证：不论$\varepsilon>0$如何小，均可找到这样的$\delta>0$，使得当$\|x\|<\delta$时算子$Q^{(x)}$将球$\|y\|\leqslant\varepsilon$变为自己．为此，我们来求出并且估计算子$Q^{(x)}$的导数．依据 1.2 中的命题 III 和 V

$$Q^{(x)\prime}(\bar{y})(y)=y-\Gamma P'_y(x, \bar{y})(y)=-\Gamma[P'_y(x, \bar{y})-P'_y(\mathbf{0}, \mathbf{0})](y),$$

由此得

$$\|Q^{(x)\prime}(\bar{y})\|\leqslant\|\Gamma\|\|P'_y(x, \bar{y})-P'_y(\mathbf{0}, \mathbf{0})\|.$$

由于假定了算子P'_Y在点$(0, 0)$的连续性，所以上述不等式右端的第二个因子可以变得任意小，通过选取充分小的ε和δ即可使得

$$\|Q^{(x)\prime}(\bar{y})\|\leqslant\alpha<1 \quad (\|x\|<\delta, \|\bar{y}\|\leqslant\varepsilon). \tag{2}$$

还要对$Q^{(x)}(0)$作出估计．注意到条件1），则

$$\|Q^{(x)}(0)\|=\|\Gamma P(x, 0)\|\leqslant\|\Gamma\|\|P(x, 0)-P(0, 0)\|,$$

而这就表明(考虑到算子P在点$(0, 0)$处的连续性)通过减小δ即可使上面那个不等式的右端变得任意小．如果必要的话，我们就认为δ已经减小到使

$$\|Q^{(x)}(0)\|\leqslant\varepsilon(1-\alpha) \qquad (\|x\|<\delta). \tag{3}$$

现在借助于有限增量公式即可很容易地证明以上所指出的论断．事实上，如果$\|x\|<\delta$, $\|y\|\leqslant\varepsilon$，则依(2)和(3)

$$\begin{aligned}\|Q^{(x)}(y)\|&\leqslant\|Q^{(x)}(0)\|+\|Q^{(x)}(y)-Q^{(x)}(0)\|\\&\leqslant\varepsilon(1-\alpha)+\sup_{0<\theta<1}\|Q^{(x)\prime}(\theta y)\|\|y\|\\&\leqslant\varepsilon(1-\alpha)+\alpha\varepsilon=\varepsilon.\end{aligned}$$

于是，算子$Q^{(x)}$将闭球$\|y\|\leqslant\varepsilon$变为自己而且对于$Q^{(x)\prime}$成立关系式(2)．这样，定理1.1的条件得以满足，因此在这个球中存在这个算子的唯一不动点$y^*=F(x)$，即

$$y^*=y^*-\Gamma P(x, y^*)$$

也就是说

$$P(x, y^*)=0.$$

算子F就是所需要的算子．事实上，a）已经验证过了是成立的．条件b）的成立可由条件1）的满足(注意这个条件可以写成$Q_0(0)=0$的形式)以及算子Q_0不动点的唯一性推知．最后，条件c)可由如下事实得知其成立：由于选取的ε是下无界的，所以它可取得任意地小．

现证F的唯一性．在球$\|y\|\leqslant\varepsilon$中算子$Q^{(x)}$（$\|x\|<\delta$）的不动

点是唯一的，由于算子 F_1 的连续性，对充分小的 η 将有

$$\|F_1(x)\| = \|F_1(x) - F_1(\mathbf{0})\| \leqslant \varepsilon \qquad (\|x\| < \eta).$$

于是，对这样的 x 就成立着等式 $F_1(x) = F(x)$.

定理 2. 如果在定理 1 中假定了算子 P 在 Ω 的每一点的连续性，则算子 F 将在 x_0 点的某个邻域中是连续的.

证. 在我们所讨论的情形下可对算子 $Q^{(x)}$ 应用定理 XVI. 1.3 的结果. 事实上，算子 $Q^{(x)}(\|x\|<\delta)$ 将球 $\|y\| \leqslant \varepsilon$ 变为自己而不等式(2)就保证了 $Q^{(x)}$ 是压缩算子(对 $\|x\|<\delta$ 的所有 $x \in \mathbf{X}$ 是一致的). 算子 $Q^{(x)}$ 对参数(在这种情况下 x 起参数的作用)的连续性可从算子 P 的连续性推知.

应用定理 XVI. 1.3，我们就得到 $y^* = F(x)$ 对 $\|x\|<\delta$ 是连续地依赖于 x 的.

注. 定理 1 和 2 具有局部的特性. 此外，在这两个定理中并未指出一种有效的用 δ 估计 ε 的方法. 这一不足将在下一章得以克服，在那里通过引用二阶导数而建立了更精确的结果.

利用下面的定理 3，在这方面可得到某些结果.

4.3. 定理 3. 设定理 1 的条件满足而且偏导数 P'_x 在 Ω 中存在并在点 (x_0, y_0) 处连续. 那么，算子 F 在 x_0 处可微，并且

$$F'(x_0) = U,$$

其中

$$U = -\Gamma P'_x(x_0, y_0) = -[P'_y(x_0, y_0)]^{-1} P'_x(x_0, y_0). \tag{4}$$

证. 如同在定理 1 的证明中那样，可以认为 $x_0 = \mathbf{0}$, $y_0 = \mathbf{0}$.

按 1.1 所述，应当证明对任意的 $\varepsilon > 0$，可找到 $\delta > 0$ 使得对任意的 $x \in \mathbf{X}(\|x\| < \delta)$ 均有

$$\|F(x) - F(\mathbf{0}) - U(x)\| \leqslant \varepsilon \|x\|,$$

由于 $F(\mathbf{0}) = \mathbf{0}$，也就是要证

$$\|F(x) - U(x)\| \leqslant \varepsilon \|x\|. \tag{5}$$

设 $F(x)=y$ 并将 U 用它的表达式(4)来代替，我们就可将(5)式左端在范数符号内的表达式写成如下形式:

$$F(x)-U(x)=y+\Gamma P'_x(0,0)x=\Gamma[P'_x(0,0)x+P'_y(0,0)y].$$

但 $P(x,y)=P(0,0)=0$, 因此按不等式(1)就有

$$\begin{aligned}\|F(x)-U(x)\| &\leqslant \|\Gamma\|\|P(x,y)-P(0,0)-P'_x(0,0)x-P'_y(0,0)y\| \\ &\leqslant \|\Gamma\|[\sup_{0<\theta,\theta_1<1}\|P'_x(\theta x,\theta_1 y)-P'_x(0,0)\|\|x\| \\ &\quad+\sup_{0<\theta,\theta_1<1}\|P'_y(\theta x,\theta_1 y)-P'_y(0,0)\|\|y\|]\leqslant\eta[\|x\|+\|y\|].\end{aligned}$$

并且, 由偏导数 P'_x 和 P'_y 的连续性可知, 只要 δ 选择得充分小, 则 η 就可变得任意小. 因此,

$$\begin{aligned}\|F(x)-U(x)\| &\leqslant\eta[\|x\|+\|F(x)\|] \\ &\leqslant\eta[\|x\|+\|U(x)\|+\|F(x)-U(x)\|].\end{aligned}$$

由此可知, 只要 η 充分小, 就有

$$\|F(x)-U(x)\|\leqslant\eta\frac{1+\|U\|}{1-\eta}\|x\|,$$

所以, 只要选择 η 使得

$$\eta\frac{1+\|U\|}{1-\eta}\leqslant\varepsilon,$$

则对所有充分小的 x, (5) 式将被满足. 这就是需要证明的.

4.4. 在最后一段里, 我们引用一个应用定理 1 的例子.

考察常微分方程组

$$y'_i(s)=f_i(s,y_k(s))\qquad(i,k=1,2,\cdots,n),\tag{6}$$

如同在 XVII. 3.3 中那样, 这个方程组可写成

$$y(s)=y(0)+\int_0^s f(t,y(t))\,dt\tag{7}$$

的形式. 假定 $f(t,0)=0$ $(t\geqslant0)$, 则方程组(6), 或者它的等价形式 (7) 有平凡解——恒等于零的解. 这个解满足初始条件 $y(0)$

$=0$.

具有重要实际价值的问题是：在什么条件下方程(7)对于充分小的初值在$[0,\infty)$上有解，而且，主要的是，当初值接近于零时能使解在整个半直线$[0,\infty)$上接近于零？下面我们给出精确的定义.

如果对每一个$\varepsilon>0$都可找到这样的$\delta>0$，使得只要$\|x\|<\delta$方程

$$y(s)=x+\int_0^s f(t,y(t))\,dt \tag{8}$$

具有唯一的有界解y^*，并且

$$\sup_{s\geqslant 0}|y^*(s)|\leqslant\varepsilon,$$

则称方程(7)的零解是稳定的．此外，如果$y^*(s)\xrightarrow[s\to\infty]{}0$，则称零解是渐近稳定的.

在考察稳定性的充分条件之前，我们先引进一个辅助性的概念.

设T是n阶方阵．我们构造一级数

$$E+\frac{1}{1!}T+\frac{1}{2!}T^2+\cdots+\frac{1}{n!}T^n+\cdots,$$

其中E是单位方阵．由于$|T^k|\leqslant|T|^k\ (k=0,1,\cdots)$，所以上面所写的级数对任何方阵$T$均收敛．它的和记为$e^T$.

如果x_0是一固定的向量，则向量函数

$$u(s)=e^{sT}x_0\qquad(s\geqslant 0)$$

满足微分方程

$$u'(s)=Tu(s), \tag{9}$$

可以如同对普通实函数一样地对此加以验证．这个事实用于下述引理的证明之中.

引理1. 设$\mu_1,\mu_2,\cdots$是矩阵T的特征值，则矩阵e^{sT}的元素

具有形如$\sum\limits_k p_k(s)e^{s\mu_k}$的形式．其中$p_k$是次数不超过$n$的多项式．

因此，若$\mu=\max\limits_k \mathrm{Re}\mu_k$，则矩阵$e^{sT}$的范数可估计如下：

$$|e^{sT}|\leqslant K_\eta\rho^{(\mu+\eta)s}, \tag{10}$$

其中η是任意一个大于零的数．

证．第一个结论可直接从常系数线性微分方程组理论的基本事实推得．这是因为，根据前面所作的说明，矩阵e^{sT}的列构成方程(9)的解的完全线性独立组，而方程(9)实质上是常系数线性微分方程组*)．

因为矩阵的范数是用矩阵诸元素平方和的平方根来估计的，所以直接可推出(10)．

现在转向对方程(7)的讨论．如同在3.3中一样，设向量函数$f(t,u)$在$(n+1)$维空间的区域G：$t\geqslant0$，$|u|\leqslant re^{-\alpha t}$，$\alpha>0$上给定，当$u=0$时关于$u$连续并且存在连续的导数$f'_u(t,u)$．我们还假定矩阵$A=f'_u(t,0)$不依赖于$t$并以$\lambda_1,\lambda_2,\cdots$表示$A$的特征值．

定理4．如果

$$\mathrm{Re}\lambda_k<0 \qquad (k=1,2,\cdots),$$

则方程(7)的零解是渐近稳定的．更确切地讲，成立如下估计式．

$$|y^*(s)|\leqslant L_x e^{-\alpha s} \qquad (s\geqslant0),$$

其中常数L_x只依赖于x并且当$x\to0$时趋于零，$\alpha>0$并且$\mathrm{Re}\lambda_k<-\alpha$　$(k=1,2,\cdots)$．

证．依据定理1，设$\mathbf{X}=\mathrm{l}_n^2$，$\mathbf{Y}=\mathbf{C}_n^\alpha$，$\mathbf{Z}=\mathbf{D}_n^\alpha$，算子$P$按如下方式确定：

$$z=P(x,y),\quad z(s)=y(s)-x-\int_0^s f(t,y(t))\,dt$$

---

*) 见 Степанов[1]214页．

并且取Ω作为这样的点$(x, y)\in \mathbf{X}\times\mathbf{Y}$ 的集合：$x\in\mathbf{X}$, $\|y\|<r$.

考虑到定理 3.3 的结果，我们立刻即可确信定理 1 的前两个条件都满足．验证条件 3）成立就比较复杂一些．首先，根据定理3.3，导数 $P'_{\mathbf{Y}}(\mathbf{0},\mathbf{0})=U$ 存在并且具有如下形式：

$$z=U(y), \qquad z(s)=y(s)-\int_0^s Ay(t)\,dt,$$

因此，需要证明算子U具有连续的逆算子．由于空间 $\mathbf{C}_n^\alpha$和$\mathbf{D}_n^\alpha$ 是完备的，所以为此只须证明算子U实现 $\mathbf{C}_n^\alpha$和 $\mathbf{D}_n^\alpha$的元素之间的一对一的对应(参见 XII.1.3)，也就是说对任何 $\tilde{z}\in\mathbf{D}_n^\alpha$ 存在唯一的元素 $\tilde{y}\in\mathbf{C}_n^\alpha$，使得 $\tilde{z}=U(\tilde{y})$．换言之，应该证明方程

$$\tilde{z}(s)=y(s)-\int_0^s Ay(t)\,dt \tag{11}$$

在 $\mathbf{C}_n^\alpha$中具有唯一的解．

这一点可象在初等情形那样，从微分方程理论的基本事实推出．所以方程(11)的解 $\tilde{y}$ 总是存在、唯一而且可表示成

$$\tilde{y}(s)=\int_0^s e^{(s-t)A}\tilde{z}'(t)\,dt+e^{sA}z(0).$$

因而，取数 σ 使得

$$\alpha<\sigma<-\max_k \mathrm{Re}\lambda_k,$$

根据估计式(10)，得到

$$\begin{aligned}
|\tilde{y}(s)| &\leqslant K\left[\int_0^s e^{-(s-t)\sigma}|\tilde{z}'(t)|\,dt+e^{-s\sigma}|\tilde{z}(0)|\right]\\
&\leqslant K\left[e^{-s\sigma}\int_0^s e^{(\sigma-\alpha)t}\|\tilde{z}\|\,dt+e^{-s\sigma}\|\tilde{z}\|\right]\\
&=K\|\tilde{z}\|e^{-s\sigma}\left[\frac{e^{(\sigma-\alpha)s}-1}{\sigma-\alpha}+1\right]\\
&=K\|\tilde{z}\|e^{-\alpha s}\left[\frac{1-e^{(\alpha-\sigma)s}}{\sigma-\alpha}+e^{(\alpha-\sigma)s}\right]\\
&\leqslant K_1\|\tilde{z}\|e^{-\alpha s},
\end{aligned}$$

因此

$$\sup|\tilde{y}(s)e^{\alpha s}|\leqslant K_1\|\tilde{z}\|<\infty. \tag{12}$$

因为

$$\tilde{y}'(s)=\tilde{z}'(x)-A\tilde{y}(s),$$

所以

$$\sup_s|\tilde{y}'(s)e^{\alpha s}|\leqslant\sup_s|\tilde{z}'(s)e^{\alpha s}|+|A|\sup_s|\tilde{y}(s)e^{\alpha s}|$$
$$\leqslant(1+K_1|A|)\|\tilde{z}\|<\infty. \tag{13}$$

将(12)与(13)联系起来即可断定 $\tilde{y}\in\mathbf{C}_n^{\alpha}$，于是定理 1 的条件完全得到了验证.

这样，根据上面的证明即可断言，只要方程(8)初始向量 x 充分小，方程 $P(x,y)=\mathbf{0}$(也就是(8))存在唯一的解 $y^*=F(x)\in\mathbf{C}_n^{\alpha}$. 其次，还有

$$|y^*(s)|\leqslant\|F(x)\|e^{-\alpha s},$$

并且，根据算子 F 的连续性，当 $x\to\mathbf{0}$ 时将有 $\|F(x)\|\to 0$.

定理证毕.

第十八章 Newton 法

本章发展了泛函方程解法的精细理论，这一理论在实方程的情形是以 Newton 法或切线法而著称．这一方法和它的某些变种是目前在实践中被应用的实际求解非线性泛函方程的不多的几个方法之一．

还应当指出的是这个方法的理论价值，因为借助于这个方法不必求解方程本身就可以作出关于方程解的存在性、唯一性以及解的分布范围等结论，而这有时比真正知道了解还重要．

本章的结果主要属于 Л. В. Канторович（参见 Канторович [9]）．叙述方法仿照 Канторович 的文章[12]．

§ 1. $P(x)=0$ 型方程

1. 1. 设 P 是将一个 Banach 空间 $\mathbf{X}$ 的开集 Ω 变到另一个 Banach 空间 $\mathbf{Y}$ 的算子．假定在 Ω 中存在算子 P 的零点，即存在使得

$$P(x^*)=0$$

的元素 x^*．

取任意元素 $x_0\in\Omega$．假定算子 P 在 Ω 中存在连续的导数，元素 $P(x_0)=P(x_0)-P(x^*)$ 可用接近于它的表达式 $P'(x_0)(x_0-x^*)$ 来代替，因而就有根据认为方程

$$P'(x_0)(x_0-x)=P(x_0)$$

的解将和 x^* 接近．但这个方程是线性方程，很容易求得它的解．这个解就是

$$x_1=x_0-[P'(x_0)]^{-1}(P(x_0))$$

(当然,要假定算子$[P'(x_0)]^{-1}$的存在).

从初始近似x_0出发继续这个过程,我们就得到序列$\{x_n\}$:

$$x_{n+1}=x_n-[P'(x_n)]^{-1}(P(x_n)) \qquad (n=0,1,\cdots). \tag{1}$$

每个x_n都是方程

$$P(x)=0 \tag{2}$$

的近似解,一般说来,n愈大则近似解就愈精确.

上面构成序列$\{x_n\}$的方法称为Newton法*).

显然,Newton法并非总能实现得了.首先,对某些n来说,x_n可能会越出Ω的范围;其次,即使不考虑这种情况,$[P'(x_n)]^{-1}$也可能不存在.

如果序列$\{x_n\}$收敛于根x^*而x_0选取得充分接近于x^*,则根据算子P'的连续性,算子$P'(x_n)$与$P'(x_0)$将相差甚小.这就给出了用简化的

$$x'_{n+1}=x_n'-[P'(x_0)]^{-1}(P(x_n')) \qquad (n=0,1,\cdots;x_0'=x_0), \tag{3}$$

去代替公式(1)的依据.一般来说,(3)式可能给出比较坏的近似解,但计算过程较之(1)要简单得多.

我们将构成序列$\{x_n'\}$的方法称为简化Newton法**).

下面,仔细地来研究Newton法(基本Newton法与简化Newton法)可实现的条件以及它的收敛性,也就是说序列$\{x_n\}$和$\{x_n'\}$趋于方程(2)的解的性质.

需要指出的是,对于方程(2)的Newton法是和应用于方程

$$x=x-\Gamma(x)(P(x)) \qquad (\Gamma(x)=[P'(x)]^{-1}) \tag{4}$$

---

*) 如果算子P是实变数的实函数,则公式(1)可写为

$$x_{n+1}=x_n-\frac{P(x_n)}{P'(x_n)} \qquad (n=0,1,\cdots),$$

这就归之于本来意义下的Newton法.

**) 导出的这两个方法常常称作Newton-Канторович法

的通常的逐步逼近法一样的. 显然, 这两个方程是等价的.

同样, 简化的牛顿法乃是对方程

$$x=x-\Gamma(x_0)(P(x)) \tag{5}$$

的逐步逼近法.

由于这种情况, 我们要先来研究一下通常的逐步逼近法.

1.2. 考察方程

$$x=S(x), \tag{6}$$

其中 S 是定义在某个 Banach 空间 $\mathbf{X}$ 中球 $\|x-x_0\|<R$ 中的算子 ($x_0\in\mathbf{X}$). 与此同时, 我们还要来考察方程

$$t=\varphi(t), \tag{7}$$

其中 φ 是定义在区间 $[t_0, t']$ ($t'=t_0+r<t_0+R$) 中的实函数. 如果

$$\left.\begin{array}{l} 1)\ \|S(x_0)-x_0\|\leqslant\varphi(t_0)-t_0; \\ 2)\ \|S'(x)\|\leqslant\varphi'(t)\ (\text{当}\|x-x_0\|\leqslant t-t_0\text{时}), \end{array}\right\} \tag{8}$$

则我们就说方程(7)控制方程(6), 或说函数 φ 控制算子 S.

定理 1. 设算子 S 在闭球 $\Omega_0(\|x-x_0\|\leqslant r)$ 中存在连续的导数, 而函数 φ 在区间 $[t_0, t']$ 中是可微的. 这时, 如果方程(7)控制方程(6), 并且方程(7)在区间 $[t_0, t']$ 中有根, 则方程(6)也有解 x^* 而且从 x_0 出发的逐步逼近序列 $\{x_n\}$:

$$x_{n+1}=S(x_n) \qquad (n=0, 1, \cdots) \tag{9}$$

收敛于 x^*.

而且,

$$\|x^*-x_0\|\leqslant t^*-t_0, \tag{10}$$

其中 t^* 表示方程(7)在 $[t_0, t_1]$ 上的最小根.

证. 首先证明方程(7)的逐步逼近

$$t_{n+1}=\varphi(t_n) \qquad (n=0, 1, \cdots) \tag{11}$$

构成收敛序列. 为此目的, 我们指出从条件 2)推知有不等式

$$\varphi'(t) \geqslant 0 \qquad (t \in [t_0, t']).$$

因此，函数 φ 在区间 $[t_0, t']$ 中是增函数．由此得知，对任何 n, t_n 都有意义，同时还有

$$t_n \leqslant \bar{t} \qquad (n=0, 1, \cdots), \tag{12}$$

其中 $\bar{t}$ 表示方程(7)的根，它的存在是在定理中作了假定的．

事实上，当 $n=0$ 时，不等式(12)是显然的．如果对于 $n=k$ 我们已经证明了(12)是成立的话，则由 φ 的单调性从 $t_k \leqslant \bar{t}$ 可得到 $\varphi(t_k) \leqslant \varphi(\bar{t})$，也就是 $t_{k+1} \leqslant \bar{t}$．按归纳法，不等式(12)对所有的 n 均成立．

同样，利用 φ 的单调性，按归纳法也可以证明序列$\{t_n\}$的单调性．

事实上，由 $t_n \leqslant t_{n+1}$ 推知 $t_{n+1} = \varphi(t_n) \leqslant \varphi(t_{n+1}) = t_{n+2}$，不等式 $t_0 \leqslant t_1$ 乃是条件 1)的推论．

这样，我们就建立了极限 $t^* = \lim\limits_{n\to\infty} t_n$ 的存在性，根据(11)以及 φ 的连续性，这个极限就是方程(7)的根．此外，从(12)可知它是方程(7)在区间$[t_0, t']$中的最小根．

现在来证明(9)的所有元素都有意义并且构成收敛序列．

首先，将(8)式写成

$$\|x_1 - x_0\| \leqslant t_1 - t_0$$

的形式，而且取 $x_1 \in \Omega_0$．设已经证明了 $x_1, x_2, \cdots, x_n \in \Omega_0$ 而且

$$\|x_{k+1} - x_k\| \leqslant t_{k+1} - t_k \qquad (k=0, 1, \cdots, n-1). \tag{13}$$

那么，按照 XVII. 1.7 中的结果，就有

$$x_{n+1} - x_n = S(x_n) - S(x_{n-1}) = \int_{x_{n-1}}^{x_n} S'(x)\,dx.$$

如果用 x 和 t 表示区间$[x_{n-1}, x_n]$和$[t_{n-1}, t_n]$中相对应的点，亦即

$$x = x_{n-1} + \tau(x_n - x_{n-1}), \quad t = t_{n-1} + \tau(t_n - t_{n-1})$$
$$(0 \leqslant \tau \leqslant 1),$$

则由(13)

$$\begin{aligned}\|x-x_0\| &\leqslant \tau\|x_n-x_{n-1}\|+\|x_{n-1}-x_{n-2}\|+\cdots+\|x_1-x_0\| \\ &\leqslant \tau(t_n-t_{n-1})+(t_{n-1}-t_{n-2})+\cdots+(t_1-t_0) \\ &= t-t_0.\end{aligned}$$

因此，由条件 2)得

$$\|S'(x)\| \leqslant \varphi'(t).$$

由此，再利用 XVII. 1.7 中的注，得到

$$\begin{aligned}\|x_{n+1}-x_n\| &= \left\|\int_{x_{n-1}}^{x_n} S'(x)\,dx\right\| \leqslant \int_{t_{n-1}}^{t_n} \varphi'(t)\,dt \\ &= \varphi(t_n)-\varphi(t_{n-1}) = t_{n+1}-t_n.\end{aligned}$$

于是，对于 $k=n$ 就证明了(13)式，同时还证得 $x_{n+1}\in\Omega_0$，这是因为

$$\begin{aligned}\|x_{n+1}-x_0\| &\leqslant \|x_{n+1}-x_n\|+\|x_n-x_{n-1}\|+\cdots+\|x_1-x_0\| \\ &\leqslant (t_{n+1}-t_n)+(t_n-t_{n-1})+\cdots+(t_1-t_0) \\ &= t_{n+1}-t_0 \leqslant t'-t_0=r.\end{aligned}$$

这样，按归纳法，对所有的 $k=0, 1, \cdots$证得了 $x_k\in\Omega_0$ 以及(13)式成立.

其次，由(13)式

$$\begin{aligned}\|x_{n+p}-x_n\| &\leqslant \|x_{n+p}-x_{n+p-1}\|+\cdots+\|x_{n+1}-x_n\| \\ &\leqslant (t_{n+p}-t_{n+p-1})+\cdots+(t_{n+1}-t_n)=t_{n+p}-t_n,\end{aligned} \tag{14}$$

故知序列$\{x_n\}$自收敛，这就表明这个序列存在极限，记之为 x^*，即 $x^*=\lim\limits_{n\to\infty}x_n$. 对(9)式取极限并考虑到算子 S 的连续性，则得到

$$x^*=S(x^*),$$

这也就是说 x^* 是方程(6)的根.

在(14)中令 $n=0$ 并让 $p\to\infty$ 取极限即得估计式(10).

注 1. 在(14)中让 $p\to\infty$ 取极限即得不等式

$$\|x^*-x_n\|\leqslant t^*-t_n \qquad (n=0,1,\cdots), \tag{15}$$

它给出了序列$\{x_n\}$收敛于x^*的速度估计.

注 2. 在定理的证明中, 条件(2) 并未在所有的范围内用到. 在证明中只要求不等式$\|S'(x)\|\leqslant\varphi'(t)$在区间$[x_{n-1}, x_n]$和$[t_{n-1}, t_n]$ $(n=1,2,\cdots)$的对应点中成立.

方程(6)在球Ω_0中可能还有另一些异于x^*的根, 甚至当控制方程(7)在区间$[t_0, t']$中只有唯一的解t^*时也可能如此. 但是却成立如下的定理.

定理 2. 设上一定理的所有条件都满足而且

$$\varphi(t')\leqslant t'.$$

这时,如果方程(7)在区间$[t_0, t']$中具有唯一的解,则方程(6)在球Ω_0中有唯一的根x^*而且从任意的初值$\tilde{x}_0\in\Omega_0$出发的逐步逼近过程收敛于该根.

证. 取初值为$\bar{t}_0=t'$而对方程(7)应用逐步逼近法. 采用与定理1完全相同的符号,可以证明序列

$$\bar{t}_{n+1}=\varphi(\bar{t}_n) \qquad (n=0,1,\cdots)$$

单调下降而下有界$(\bar{t}_n\geqslant t^*)$, 因此具有极限$\bar{t}$, 而且$\bar{t}$作为方程(7)的根应与t^*相等.

现在证明:对于方程(6)的逐步逼近过程对以任何元素$\tilde{x}_0\in\Omega_0$为初值均收敛,因而就给出了方程(6)的根.

逐步逼近具有如下形式

$$\tilde{x}_{n+1}=S(\tilde{x}_n) \qquad (n=0,1,\cdots).$$

由于

$$\tilde{x}_1-x_1=S(\tilde{x}_0)-S(x_0)=\int_{x_0}^{\tilde{x}_0}S'(x)\,dx$$

再次应用 XVII.1.7 的注,我们就得到

$$\|\tilde{x}_1-x_1\|\leqslant\int_{t_0}^{\bar{t}_0}\varphi'(t)\,dt=\varphi(\bar{t}_0)-\varphi(t_0)=\bar{t}_1-t_1,$$

由此易见

$$\|\tilde{x}_1-x_0\|\leqslant\|\tilde{x}_1-x_1\|+\|x_1-x_0\|$$
$$\leqslant(\bar{t}_1-t_1)+(t_1-t_0)=\bar{t}_1-t_0\leqslant r,$$

这表明 $\tilde{x}\in\Omega_0$.

设已经证得了

$$\tilde{x}_k\in\Omega_0,\ \|\tilde{x}_k-x_k\|\leqslant\bar{t}_k-t_k,\ (k=1,2,\cdots,n).\qquad(16)$$

那么

$$\tilde{x}_{n+1}-x_{n+1}=S(\tilde{x}_n)-S(x_n)=\int_{x_n}^{\tilde{x}_n}S'(x)dx.$$

但如果 x 和 t 是区间 $[x_n,\tilde{x}_n]$ 和 $[t_n,\bar{t}_n]$ 的对应点，即

$$x=x_n+\tau(\tilde{x}_n-x_n),\quad t=t_n+\tau(\bar{t}_n-t_n)\quad(0\leqslant\tau\leqslant1),$$

则

$$\|x-x_0\|\leqslant\tau\|\tilde{x}_n-x_n\|+\|x_n-x_{n-1}\|+\cdots+\|x_1-x_0\|$$
$$\leqslant\tau(\bar{t}_n-t_n)+(t_n-t_{n-1})+\cdots+(t_1-t_0)$$
$$=t-t_0.$$

因此，对于所指出的 x 和 t 将有 $\|S'(x)\|\leqslant\varphi'(t)$，而且

$$\|\tilde{x}_{n+1}-x_{n+1}\|=\int_{t_n}^{\bar{t}_n}\varphi'(t)dt=\varphi(\bar{t}_n)-\varphi(t_n)=\bar{t}_{n+1}-t_{n+1},$$

于是

$$\|\tilde{x}_{n+1}-x_0\|\leqslant\|\tilde{x}_{n+1}-x_{n+1}\|+\|x_{n+1}-x_0\|$$
$$\leqslant(\bar{t}_{n+1}-t_{n+1})+(t_{n+1}-t_0)=\bar{t}_{n+1}-t_0\leqslant r,$$

这表明 $\tilde{x}_{n+1}\in\Omega_0$.

根据归纳法，可断言(16)式对所有的 $k=1,2,\cdots$ 均成立.

由于序列 $\{t_n\}$ 和 $\{t_n'\}$ 具有共同的极限 t^*，所以，根据(16)，从序列 $\{x_n\}$ 的收敛性可推出序列 $\{\tilde{x}_n\}$ 的收敛性，并且有等式：

$$\lim_{n\to\infty}\tilde{x}_n=\lim_{n\to\infty}x_n=x^*.\qquad(17)$$

于是，就证明了不论初始逼近 $\tilde{x}_0\in\Omega_0$ 如何选取，逐步逼近过

程均收敛于 x^*. 由此可立即推出方程(6)解的唯一性. 事实上，如果 $\tilde{x}\in\Omega_0$ 是这个方程的根，那么，取 $\tilde{x}_0=\tilde{x}$, 我们显然得到对任何 $n=1,2,\cdots$, 均有

$$\tilde{x}_n=\tilde{x},$$

由(17)知 $\tilde{x}=x^*$, 这就是所要证明的.

1.3. 现在转向对方程(2)的 Newton 法的研究.

考察方程(2)的同时我们还考察实方程

$$\psi(t)=0. \tag{18}$$

假定算子 P 定义在空间 $\mathbf{X}$ 的球 $\Omega(\|x-x_0\|<R)$ 内，而且在闭球 $\Omega_0(\|x-x_0\|\leqslant r)$ 内具有连续的二阶导数. 至于函数 ψ, 将假定在区间 $[t_0, t']$ $(t'=t_0+r)$ 中是二次连续可微的.

借助于定理 1 容易得到关于简化 Newton 法收敛性的下述结果.

定理 3. 设如下诸条件满足:

1) 存在连续线性算子 $\Gamma_0=[P'(x_0)]^{-1}$;

2) $c_0=-\dfrac{1}{\psi'(t_0)}>0$;

3) $\|\Gamma_0(P(x_0))\|\leqslant c_0\psi(t_0)$;

4) $\|\Gamma_0 P''(x)\|\leqslant c_0\psi''(t)$, 如果 $\|x-x_0\|\leqslant t-t_0\leqslant r$;

5) 方程(18)在区间 $[t_0, t']$ 中存在根 $\bar{t}$.

这时对方程(2)和(18)的(分别以 x_0 和 t_0 为初值的)简化 Newton 过程分别收敛于这两个方程的解 x^* 和 t^*, 而且

$$\|x^*-x_0\|\leqslant t^*-t_0. \tag{19}$$

证. 上面已经指出. 对于方程(2)的简化 Newton 法是和对于方程(5)也就是对于方程

$$x=S(x) \qquad (S(x)=x-\Gamma_0(P(x))) \tag{20}$$

的逐步逼近法一样的. 同样，对于方程(18)的简化 Newton 过程

也可代之以对方程

$$t=\varphi(t) \qquad (\varphi(t)=t+c_0\psi(t)) \tag{21}$$

的逐步逼近过程. 我们来证明对于方程(20)和(21)定理 1 的条件是满足的.

由(20)和(21)得到

$$S(x_0)-x_0=-\Gamma_0(P(x_0)),$$
$$\varphi(t_0)-t_0=c_0\psi(t_0).$$

因此, 从条件 3)

$$\|S(x_0)-x_0\|\leqslant\varphi(t_0)-t_0.$$

其次, 利用微分法则(XVII. 1. 2), 得到

$$S'(x)=I-\Gamma_0P'(x), S''(x)=-\Gamma_0P''(x).$$

因此

$$S'(x)=S'(x)-S'(x_0)=\int_{x_0}^{x} S''(x)dx=-\int_{x_0}^{x} \Gamma_0P''(x)dx.$$

因而, 若取 x 和 t 使得 $\|x-x_0\|\leqslant t-t_0$, 则根据 XVII. 1. 7 的注和条件 4), 得到

$$\|S'(x)\|\leqslant\int_{t_0}^{t} c_0\psi''(\tau)d\tau=c_0\psi'(t)-c_0\psi'(t_0)$$
$$=1+c_0\psi'(t_0)=\varphi'(t).$$

这样, 可见方程(21)是(20)的控制方程. 对这两个方程应用定理 1 就推出了所要证明的结果.

注. 正象在定理 1 中那样, 序列$\{x_n'\}$(见(3))收敛于 x^* 的速度由不等式

$$\|x^*-x_n'\|\leqslant t^*-t_n' \ (n=0, 1, \cdots; x_0'=x_0; t_0'=t_0) \tag{22}$$

来估计, 其中 t_n' 是对方程(18)的简化 Newton 法的逐步逼近.

应用定理 2 到方程(20)和(21)可立即得到方程(2)的解的唯一性.

定理 4. 设定理 3 的条件全保持而且

$$\psi'(t)\leqslant 0. \tag{23}$$

那么，当方程(18)在$[t_0, t']$中有唯一的根时，方程(2)在球 Ω_0 中仅有一个解.

要证明这个定理只须指出由(23)可推出

$$\varphi(t')=t'+c_0\psi(t')\leqslant t',$$

也就是说定理 2 的补充条件是满足的.

注. 因为 t^* 是方程(18)的最小根，若令 $t'=t^*$，则可应用唯一性定理. 这样，就保证方程(2)的解 x^* 在任何情形下于球

$$\|x-x_0\|\leqslant t^*-t_0$$

中的唯一性.

1.4. 基本 Newton 过程的收敛性可根据关于简化过程收敛性的定理 3 来确立.

定理 5. 设定理 3 的所有条件都满足，则对方程(2)的以 x_0 为初值的基本 Newton 过程产生出收敛于方程(2)的根 x^* 的序列 $\{x_n\}$.

证. 简化 Newton 过程和基本 Newton 过程的第一步是一样的，因此 x_1 有意义而且 $x_1\in\Omega_0$. 我们现在来验证：如果用 x_1 和 t_1 分别代替定理 3 中的 x_0 和 t_0，则该定理的所有条件都不被破坏. 仍然使用定理 3 的符号，考察算子

$$I-\Gamma_0P'(x_1)=-\Gamma_0(P'(x_1)-P'(x_0))=-\int_{x_0}^{x_1}\Gamma_0P''(x)\,dx.$$

利用 XVII.1.7 中的注，得到

$$\|I-\Gamma_0P'(x_1)\|\leqslant\int_{t_0}^{t_1}c_0\psi''(t)\,dt=1+c_0\psi'(t_1)=q.$$

因为 $\psi''(t)\geqslant 0$ 在整个区间$[t_0, t']$上成立，而 $\psi(t_0)\geqslant 0$，所以函数 $\psi(t)$的极小值不能在点 t^*的左端点达到. 由于 $t_1\leqslant t^*$以及 $\psi'(t_0)$

<0，则 $\psi'(t_1)<0$*)，这表明 $q<1$.

根据 Banach 定理(参见 V.4.5)，存在连续线性算子

$$U=[\Gamma_0 P'(x_1)]^{-1}\in B(\mathbf{X},\mathbf{Y}),$$

并且

$$\|U\|\leqslant\frac{1}{1-q}=\frac{\psi(t_0)}{\psi'(t_1)}=\frac{c_0}{c_1}\qquad\left(c_1=-\frac{1}{\psi'(t_1)}\right).\qquad(24)$$

因此，存在连续线性算子

$$\Gamma_1=[P'(x_1)]^{-1}=U\Gamma_0,$$

这就验证了定理 3 的条件 1)和 2).

现在来证明

$$\|\Gamma_0(P(x_1))\|\leqslant c_0\psi(t_1).\qquad(25)$$

根据 Taylor 公式(XVII.2.5):

$$\begin{aligned}\Gamma_0(P(x_1))&=\Gamma_0(P(x_0))+\Gamma_0P'(x_0)(x_1-x_0)\\&\quad+\int_{x_0}^{x_1}\Gamma_0P''(x)(x-x_1,\cdot)dx\\&=(x_0-x_1)+(x_1-x_0)+\int_{x_0}^{x_1}\Gamma_0P''(x)(x_1-x,\cdot)dx\\&=\int_{x_0}^{x_1}\Gamma_0P''(x)(x_1-x,\cdot)dx.\end{aligned}$$

类似地，

$$c_0\psi(t_1)=c_0\int_{t_0}^{t_1}\psi''(t)(t_1-t)\,dt.$$

因为对区间$[x_0,x_1]$和$[t_0,t_1]$的对应点 x 和 t 有

$$\begin{aligned}\|\Gamma_0P''(x)(x_1-x,\cdot)\|&\leqslant\|\Gamma_0P''(x)\|\|x_1-x\|\\&\leqslant c_0\psi''(t)(t_1-t),\end{aligned}$$

---

*) 如果 $\psi'(t_1)=0$，则 $t_1=t^*$ 并且 $\psi(t_1)=0$，于是

$$0=\psi(t_1)-\psi(t_0)-\psi'(t_0)(t_1-t_0)=\int_{t_0}^{t_1}\psi''(t)(t_1-t)dt.$$

这只有当 $\psi''(t)=0$ 在区间$[t_0,t_1]$上成立时才有可能.而由此得 $\psi'(t_0)=\psi'(t_1)=0$.

所以

$$\left\|\int_{x_0}^{x_1} \Gamma_0 P''(x)(x_1-x,\cdot)dx\right\| \leqslant \int_{t_0}^{t_1} c_0\psi''(t)(t_1-t)dt,$$

这正是(25).

注意到(25)和(24), 我们就得到

$$\|\Gamma_1(P(x_1))\| = \|U\Gamma_0(P(x_1))\| \leqslant \|U\| \|\Gamma_0(P(x_1))\|$$

$$\leqslant \frac{c_1}{c_0} c_0\psi(t_1) = c_1\psi(t_1),$$

亦即条件 3)同样满足.

条件 4)可用同样的方法来验证. 首先指出: 如果$\|x-x_1\| \leqslant t-t_1$, 则更有$\|x-x_0\| \leqslant t-t_0$. 因此

$$\|\Gamma_1 P''(x)\| = \|U\Gamma_0 P''(x)\| \leqslant \|U\| \|\Gamma_0 P''(x)\|$$

$$\leqslant \frac{c_1}{c_0} c_0\psi''(t) = c_1\psi''(t).$$

最后来说明条件 5)也未被破坏. 这是因为根 $\bar{t}$ 位于区间$[t^*, t']$之中, 当然就更在更大的区间$[t_1, t']$中.

进行类似的讨论, 也就是说如果将 x_1 和 t_1 分别变到 x_2 和 t_2 等等, 我们就会证明定理 3 的条件亦不被破坏.

于是, 所有的 x_n 均有意义, 这就是说已经证明了 Newton 过程是能实施的.

因为序列$\{t_n\}$显然是上升且有界的, 所以它有极限 $\tilde{t}$, 而由不等式

$$\|x_{n+1}-x_n\| \leqslant t_{n+1}-t_n \qquad (n=0, 1, \cdots)$$

可知序列$\{x_n\}$同样存在极限 $\tilde{x} = \lim\limits_{n\to\infty} x_n$.

由(1)得到关系式

$$P(x_n) + P'(x_n)(x_{n+1}-x_n) = 0 \quad (n=0, 1, \cdots). \qquad (26)$$

对任意的 $x \in \Omega_0$, 根据有限增量公式, 有

$$\Gamma_0[P'(x)-P'(x_0)] \leqslant$$

$$\leqslant \|x-x_0\| \sup_{0<\theta<1} \|\Gamma_0 P''(x_0+\theta(x-x_0))\|$$

$$\leqslant r \max_{[t_0, t']} c_0\psi''(t),$$

由此推知$\|P'(x_n)\|$是总体有界的. 对(26)取极限,得

$$P(\tilde{x})=\mathbf{0},$$

也就是说 $\tilde{x}$ 是方程(2)的根.

用同样的方法可证明 $\tilde{t}$ 是方程(18)的根,并且因 $\tilde{t}\leqslant t^*$,而 t^* 是最小的根,所以 $\tilde{t}=t^*$.

从明显的关系式$\|\tilde{x}-x_0\|\leqslant\tilde{t}-t_0=t^*-t_0$以及定理 4 的注可知 $\tilde{x}=x^*$.

定理完全得证.

注. 序列$\{x_n\}$收敛于 x^* 的速度由下述不等式

$$\|x^*-x_n\|\leqslant t^*-t_n \qquad (n=0,1,\cdots) \tag{27}$$

给出估计.

这一估计可从不等式(22)推出, 因为 x_n 和 t_n 可看成分别以 x_{n-1} 和 t_{n-1} 为初值的简化 Newton 法的首次逼近.

1.5. 定理 3 或定理 5 通常是难以直接利用的,这是因为其中含有一个不确定的函数 ψ 的缘故. 因此, 有实际意义的是如下的定理.

定理 6. 如同以前一样, 设算子 P 定义在 Ω 上而且在 Ω_0 中具有连续的二阶导数. 此外, 还设

1) 存在连续算子 $\Gamma_0=[P'(x_0)]^{-1}$;

2) $\|\Gamma_0(P(x_0))\|\leqslant\eta$;

3) $\|\Gamma_0 P''(x)\|\leqslant K \qquad (x\in\Omega_0)$.

这时, 若

$$h=K\eta\leqslant\frac{1}{2} \tag{28}$$

以及

$$r \geqslant r_0 = \frac{1-\sqrt{1-2h}}{h}\eta, \tag{29}$$

则方程(2)具有解 x^* 并且 Newton 过程(基本的和简化的)收敛于该解. 同时

$$\|x^* - x_0\| \leqslant r_0. \tag{30}$$

此外，如果当 $h<\frac{1}{2}$ 时

$$r < r_1 = \frac{1+\sqrt{1-2h}}{h}\eta, \tag{31}$$

而当 $h=\frac{1}{2}$ 时

$$r \leqslant r_1 \tag{32}$$

的话，则解 x^* 在球 Ω_0 中是唯一的.

基本 Newton 过程的收敛速度由不等式

$$\|x^* - x_n\| \leqslant \frac{1}{2^n}(2h)^{2^n}\frac{\eta}{h} \qquad (n=0, 1, \cdots) \tag{33}$$

来刻划，而简化 Newton 过程$\left(h<\frac{1}{2}\right)$的收敛速度则由

$$\|x^* - x'_n\| \leqslant \frac{\eta}{h}(1-\sqrt{1-2h})^{n+1} \qquad (n=0, 1, \cdots) \tag{34}$$

刻划.

证. 考察区间$[0, r]$上的实函数

$$\psi(t) = Kt^2 - 2t + 2\eta = Kt^2 - 2t + \frac{2h}{K} = \frac{h}{\eta}t^2 - 2t + 2\eta.$$

今证算子 P 和函数 ψ 满足定理 3 的所有条件(而这就表明定理 5 的条件也满足). 事实上，条件 1)—4)是显然满足的；其次，因为方程

$$\psi(t) = 0 \tag{35}$$

的根是

$$r_0=\frac{1-\sqrt{1-2h}}{h}\eta,\quad r_1=\frac{1+\sqrt{1-2h}}{h}\eta,$$

所以(28)式就保证了根是实的，而条件(29)则保证了最小根 r_0 包含在区间$[0, r]$中. 由于 $t^*=r_0$，所以不等式(30)也就是定理3的不等式(19).

解的唯一性的确定乃是定理4的推论，因为，对不论怎样的 $h\leqslant\frac{1}{2}$，在定理的条件下 $\psi(r)\leqslant 0$，而且方程(35)的根在区间$[0, r]$中是唯一的.

现在来建立刻划 Newton 过程收敛速度的不等式(33)和(34). 注意定理3和定理5的注，则只要考察关于方程(35)的基本 Newton 过程和简化 Newton 过程的逐次逼近的实数序列$\{t_n\}$与$\{t_n'\}$就够了.

先来看基本 Newton 法. 记

$$c_n=-\frac{1}{\psi'(t_n)},\quad \eta_n=c_n\psi(t_n),$$

$$K_n=c_n\psi''(t_n)=2Kc_n,\quad \eta=K_n\eta_n$$

并且用下标少1的这些量来表示 η_n, K_n, h_n. 为此，注意有

$$t_{k+1}-t_k=-\frac{\psi(t_k)}{\psi'(t_k)}=\eta_k\qquad(k=0, 1, \cdots).\tag{36}$$

因而，按照二次多项式的 Taylor 公式

$$\begin{aligned}\eta_n&=c_n\psi(t_n)=c_n\psi(t_{n-1}+\eta_{n-1})\\&=c_n\left[\frac{1}{2}\psi''(t_{n-1})\eta_{n-1}^2+\psi'(t_{n-1})\eta_{n-1}+\psi(t_{n-1})\right]\\&=c_n\left[K\eta_{n-1}^2-\frac{\eta_{n-1}}{c_{n-1}}+\frac{\eta_{n-1}}{c_{n-1}}\right]\\&=c_nK\eta_{n-1}^2=\frac{1}{2}\frac{c_n}{c_{n-1}}2Kc_{n-1}\eta_{n-1}^2=\end{aligned}$$

$$=\frac{1}{2}\frac{c_n}{c_{n-1}}K_{n-1}\eta_{n-1}^2.$$

但

$$\frac{c_{n-1}}{c_n}=\frac{\psi'(t_n)}{\psi'(t_{n-1})}=\frac{\psi'(t_{n-1})+\psi''(t_{n-1})\eta_{n-1}}{\psi'(t_{n-1})}$$

$$=1-K_{n-1}\eta_{n-1}=1-h_{n-1}, \tag{37}$$

由此得到

$$\eta_n=\frac{1}{2}\frac{K_n\eta_{n-1}^2}{1-h_{n-1}}=\frac{\eta_{n-1}}{2}\cdot\frac{h_{n-1}}{1-h_{n-1}}. \tag{38}$$

借助于(37)类似地可得

$$K_n=2c_nK=2Kc_{n-1}\frac{c_n}{c_{n-1}}=\frac{K_{n-1}}{1-h_{n-1}}.$$

由此,

$$h_n=K_n\eta_n=\frac{1}{2}\frac{K_{n-1}\eta_{n-1}h_{n-1}}{(1-h_{n-1})^2}=\frac{1}{2}\left[\frac{h_{n-1}}{1-h_{n-1}}\right]^2. \tag{39}$$

由(38)和(39)并注意 $h_n\leqslant\frac{1}{2}$, 我们就得到估计

$$\eta_n\leqslant h_{n-1}\eta_{n-1},\quad h_n\leqslant 2h_{n-1}^2\quad (n=1, 2, \cdots). \tag{40}$$

因而, $h_n\leqslant\frac{1}{2}[2h_0]^{2^n}=\frac{1}{2}[2h]^{2^n}$ 并且

$$\eta_n\leqslant h_{n-1}\eta_{n-1}\leqslant h_{n-1}h_{n-2}\eta_{n-2}$$

$$\leqslant\cdots\leqslant h_{n-1}h_{n-2}\cdots h_0\eta_0\leqslant\frac{1}{2}[2h]^{2^n-1}\eta.$$

由此, 注意到(36), 我们就得到

$$t^*-t_n=(t_{n+1}-t_n)+(t_{n+2}-t_{n+1})+\cdots$$

$$=\eta_n+\eta_{n+1}+\cdots$$

$$\leqslant\frac{1}{2^n}[2h]^{2^n-1}\eta\left\{1+\frac{1}{2}[2h]^{2^n}+\frac{1}{2^2}[2h]^{2^{n+1}}+\cdots\right\}$$

$$\leqslant\frac{2}{2^n}[2h]^{2^n-1}\eta=\frac{1}{2^n}[2h]^{2^n}\frac{\eta}{h}.$$

由于(27)，这就给出了所要的估计(33).

现在来看简化 Newton 法的误差估计. 假定 $h<\frac{1}{2}$. φ 仍然表示定理 3 的证明中那同一个函数，则可写

$$t^*-t'_n=\varphi(t^*)-\varphi(t'_{n-1})=\varphi'(\tilde{t}_n)(t^*-t'_{n-1}),$$

其中

$$\tilde{t}_n=\frac{t^*+t'_{n-1}}{2}.$$

但

$$\varphi'(t)=1+c_0\psi'(t)=Kt,$$

因而

$$\varphi'(\tilde{t}_n)=K\tilde{t}_n\leqslant Kt^*=1-\sqrt{1-2h}.$$

由此得到

$$t^*-t_n'\leqslant[1-\sqrt{1-2h}](t^*-t'_{n-1}).$$

用同样的方法估计 $t^*-t'_{n-1}$ 并继续此过程，最后得到

$$t^*-t'_n\leqslant[1-\sqrt{1-2h}]^n(t^*-t'_0)=\frac{\eta}{h}[1-\sqrt{1-2h}]^{n+1},$$

将此式代入(22)，即得

$$\|x^*-x'_n\|\leqslant\frac{\eta}{h}[1-\sqrt{1-2h}]^{n+1},$$

这正是不等式(34).

定理全部得证.

注 1. 定理的条件 2)和 3)可用条件

1′) $\|\Gamma_0\|\leqslant B'$;

2′) $\|P(x_0)\|\leqslant\eta'$;

3′) $\|P''(x_0)\|\leqslant K'\quad(x\in\Omega_0)$

来代替. 这时 h, r_0 和 r_1 分别是

$$h=K'B'^2\eta',$$

$$r_0 = \frac{1-\sqrt{1-2h}}{h} B'\eta',$$

$$r_1 = \frac{1+\sqrt{1-2h}}{h} B'\eta'.$$

注 2. 条件(28), (29)和(31)(或(32))的意义就在于: 破坏了(28), (29)中任何一个条件都将使二次方程(35)没有解, 而要是没有条件(31)或(32)则唯一性不成立*).

§ 2. Newton 法收敛性定理的推论

本节指出 Newton 法收敛性定理(定理 1.6)的某些推论. 和上一节一样, 我们将考察方程

$$P(x)=\mathbf{0}, \tag{1}$$

并且关于算子 P 的定义域, 可微性等均保持上一节的假定和记法.

2.1. 首先指出定理 1.6 的下述推广, 即不须要假定逆算子 $\Gamma_0=[P'(x_0)]^{-1}$ 存在而只要求存在一个与 Γ_0 接近的算子.

定理 1. 设存在具有连续逆的线性算子 $\Gamma\in B(\mathbf{Y},\mathbf{X})$ 并且满足条件:

1) $\|\Gamma(P(x_0))\|\leqslant\bar{\eta}$;

2) $\|\Gamma P'(x_0)-I\|\leqslant\delta$;

3) $\|\Gamma P''(x)\|\leqslant\bar{K}\ (x\in\Omega_0)$.

如果

$$\bar{h}=\frac{\bar{K}\bar{\eta}}{(1-\delta)^2}\leqslant\frac{1}{2}, \quad \delta<1 \tag{2}$$

并且

$$r\geqslant\bar{r}_0=\frac{1-\sqrt{1-2\bar{h}}}{\bar{h}}\ \frac{\bar{\eta}}{1-\delta}, \tag{3}$$

*) 在条件(32) 之下, 方程 (35) 有唯一的解, 但方程 $\psi(t)-\varepsilon=0(\varepsilon>0)$ 在区间$[0,r]$中就有两个解了.

则方程(1)存在解 $x^*\in\Omega_0$. 此外, 如果

$$
\left.\begin{array}{ll}
\text{当 } \bar{h}<\frac{1}{2} \text{ 时} & r<\bar{r}_1 \\
\text{当 } \bar{h}=\frac{1}{2} \text{ 时} & r\leqslant\bar{r}_1
\end{array}\right.
\quad\left(\bar{r}_1=\frac{1+\sqrt{1-2\bar{h}}}{\bar{h}}\,\frac{\bar{\eta}}{1-\delta}\right),
$$

则解是唯一的.

这时, 若 $\{x_n\}$ 是基本 Newton 过程的逼近序列, 则

$$
\|x^*-x_n\|\leqslant\frac{1}{2^n}[2\bar{h}]^{2^n}\frac{\bar{\eta}}{\bar{h}(1-\delta)}\qquad(n=0,1,\cdots).
$$

证. 我们来验证定理 1.6 的所有条件均依然满足. 由第二个条件推知存在连续线性算子

$$
U=[\Gamma P'(x_0)]^{-1},
$$

并且

$$
\|U\|\leqslant\frac{1}{1-\delta}.
$$

由此推知存在连续线性算子

$$
\Gamma_0=[P'(x_0)]^{-1}=U\Gamma.
$$

此外

$$
\|\Gamma_0(P(x_0))\|\leqslant\|U\|\,\|\Gamma(P(x_0))\|\leqslant\frac{\eta}{1-\delta},
$$

$$
\|\Gamma_0P''(x)\|\leqslant\|U\|\,\|\Gamma P''(x)\|\leqslant\frac{\bar{K}}{1-\delta}\qquad(x\in\Omega_0).
$$

因此, 只要在定理 1.6 中用 $\frac{\bar{\eta}}{1-\delta}$ 代替 η 而用 $\frac{\bar{K}}{1-\delta}$ 代替 K 就行了.

2.2. 考察与简化的 Newton 法类似的逐步逼近过程, 在这个过程中是用一个接近于 $\Gamma_0=[P'(x_0)]^{-1}$ 的算子 Γ 来代替算子 Γ_0. 换言之, 构成序列 $\{\tilde{x}_n\}$:

$$
\tilde{x}_{n+1}=\tilde{x}_n-\Gamma(P(\tilde{x}_n))\qquad(n=0,1,\cdots;\ \tilde{x}_0=x_0).\tag{4}
$$

定理 2. 在定理 1 的条件下, 过程(4)是可实现的, 亦即 $\tilde{x}_n\in\Omega_0$

$(n=0, 1, \cdots)$ 而且

$$\lim_{n\to\infty}\tilde{x}_n = x^*.$$

证. 和在定理 1.3 中一样，在证明简化 Newton 过程的收敛性时要利用定理 1.1, 在这个定理中设

$$S(x) = x - \Gamma(P(x)) \qquad (x \in \Omega_0)$$

以及

$$\varphi(t) = \frac{1}{2}\overline{K}t^2 + \delta t + \overline{\eta} = t + \psi(t)$$

$$(\psi(t) = \frac{1}{2}Kt^2 - (1-\delta)t + \overline{\eta}, t \in [t_0, t'], t_0 = 0, t' = r).$$

于是

$$\|S(x_0) - x_0\| = \|-\Gamma(P(x_0))\| \leqslant \overline{\eta} = \varphi(t_0) - t_0.$$

其次, 如果 $\|x - x_0\| \leqslant t \leqslant r$, 则

$$\begin{aligned}\|S'(x)\| &= \|I - \Gamma P'(x)\| \\ &\leqslant \|I - \Gamma P'(x_0)\| + \|\Gamma(P'(x_0) - P'(x))\| \\ &\leqslant \delta + \left\|\int_{x_0}^{x} \Gamma P''(y)dy\right\| \leqslant \delta + \int_0^t K d\tau \\ &= \delta + Kt = \varphi'(t).\end{aligned}$$

因此, 方程

$$x = S(x) \tag{5}$$

以方程 $t = \varphi(t)$ 为控制方程, 后一方程的解是

$$t^* = \frac{1 - \sqrt{1 - 2\overline{h}}}{\overline{h}} \frac{\overline{\eta}}{1 - \delta} = \overline{r}_0 \leqslant r.$$

从而可知方程(5)有解，这个解应当就是 x^*(因为方程(1)的解 x^* 在球 $\|x - x_0\| \leqslant \overline{r}_0$ 中是唯一的).

注. 关于序列 $\{\tilde{x}_n\}$ 收敛于 x^* 的速度估计，可完全类似地按照在 1.4 中对简化 Newton 过程的处理办法推导出来. 也就是说,

我们建议读者去证明：当 $h<\frac{1}{2}$ 时

$$\|x^*-\tilde{x}_n\|\leqslant\frac{1}{\bar{h}}[1-(1-\delta)\sqrt{1-2\bar{h}}]^{n+1}\frac{\bar{\eta}}{(1-\delta)^2}$$

$$(n=0,1,\cdots).$$

2.3. 设定理 1.6 的条件对于 $h<1/2$ 成立. 我们现在指出 Newton 过程具有熟知的稳定性，也就是说，如果取的初值不是 x_0 而是任意一个充分接近于 x_0 的元素 $x_0'\in\Omega_0$，则 Newton 过程的收敛性不被破坏.

定理 3. 设定理 1.6 的条件满足且其中常数为 η, K 和 $h=K\eta<\frac{1}{2}$. 这时，如果

$$\|x_0'-x_0\|\leqslant\varepsilon\quad\left(\varepsilon=\frac{1-2h}{4K}=\frac{1-2K\eta}{4K}\right),\tag{6}$$

则以 x_0' 为初始值的基本 Newton 过程和简化 Newton 过程均收敛.

证. 利用定理 1，取其中的 Γ 为 $\Gamma_0=[P'(x_0)]^{-1}$，取 x_0 为 x_0'. 这时的条件 1）就是

$$\begin{aligned}&\|\Gamma_0(P(x_0'))\|\\&\quad=\|\Gamma_0[P(x_0)+P'(x_0)(x_0'-x_0)\\&\qquad\qquad+\int_{x_0}^{x_0'}P''(x)(x_0'-x,\cdot)dx]\|\\&\quad\leqslant\eta+\varepsilon+K\frac{\varepsilon^2}{2},\end{aligned}$$

因此，作为 $\bar{\eta}$ 应取 $\eta+\varepsilon+K\frac{\varepsilon^2}{2}$.

$$\begin{aligned}\|\Gamma_0P'(x_0')-I\|&=\|\Gamma_0[P'(x_0')-P'(x_0)]\|\\&=\|\int_{x_0}^{x_0'}\Gamma_0P''(x)dx\|.\end{aligned}$$

这样一来，$\bar{\delta}=K\varepsilon$。

最后，$\bar{K}=K$。

这时，$\bar{h}$ 为

$$\bar{h}=\frac{K\bar{\eta}}{(1-\bar{\delta})^2}=\frac{K\left(\eta+\varepsilon+\frac{K}{2}\varepsilon^2\right)}{(1-K\varepsilon)^2}=\frac{1}{2}\,\frac{K^2\varepsilon^2+2K\varepsilon+2K\eta}{K^2\varepsilon^2-2K\varepsilon+1}=\frac{1}{2}.$$

定理得证.

注. 如果 $h\geqslant 4\sqrt{2}-5.5\approx 0.16$，则球 $\|x-x_0\|\leqslant\varepsilon$ 包含在球 $\|x-x_0\|\leqslant r_0$ 中，因而也包含在球 Ω_0 中．要是 $h<4\sqrt{2}-5.5$ 的话，则一般说来，条件(6)就已经不能保证 $x_0'\in\Omega_0$ 了，因此这时就应补充 $x_0'\in\Omega_0$ 这一条件.

2.4. 在参与估计初始逼近 x_0 与解 x^* 的接近程度（亦即估计值 r_0 时）的很多场合都出现过算子 Γ_0．这个算子的意义就在于使得能将解的分布区域精确化．事实上，知道了 Γ_0，我们就可以求得牛顿法的下一逼近 x_1 并将它应用于定理 1 中.

定理 4. 设存在连续线性算子 $\Gamma_0=[P'(x_0)]^{-1}$ 并且满足条件：

1) $\|\Gamma_0(P(x_0))\|\leqslant\eta$;

2) $\|\Gamma_0(P(x_1))\|\leqslant\eta_1\qquad(x_1=x_0-\Gamma_0(P(x_0)))$;

3) $\|\Gamma_0P''(x)\|\leqslant K\qquad(x\in\Omega_0)$;

4) $h_1=\dfrac{K\eta_1}{(1-K\eta)^2}\leqslant\dfrac{1}{2}$.

那么，方程(1)存在解 x^*，并且

$$\|x^*-x_1\|\leqslant\frac{1-\sqrt{1-2h_1}}{h_1}\,\frac{\eta_1}{1-K\eta}. \tag{7}$$

证. 我们来验证定理 1 的各条件．在定理 1 中取 $\Gamma=\Gamma_0$ 并且用 x_1 代替 x_0，于是

$$\|\Gamma_0(P(x_1))\|\leqslant\eta_1,$$

$$\|\Gamma_0 P'(x_1)-I\|\leqslant\|\Gamma_0[P'(x_1)-P'(x_0)]\|$$
$$=\left\|\int_{x_0}^{x_1}\Gamma_0 P''(x)\,dx\right\|\leqslant K\|x_1-x_0\|\leqslant K\eta.$$
所以，作为 $\bar{\eta}$, δ 和 $\bar{h}$ 应分别取 η_1, $K\eta$ 和 h_1.

按照定理 1，解 x^* 位于球 $\|x-x_1\|\leqslant\bar{r}_0$ 中. 因为这时
$$\bar{r}_0=\frac{1-\sqrt{1-2\bar{h}}}{\bar{h}}\,\frac{\bar{\eta}}{1-\delta}=\frac{1-\sqrt{1-2h_1}}{h_1}\,\frac{\eta_1}{1-K\eta},$$
由此得到估计式(7).

注 1. 可能会发生这种情况，即虽然 $h=K\eta>\frac{1}{2}$，但 $h_1\leqslant\frac{1}{2}$. 这时虽然不能再应用定理 1.6，但定理 4 却依然能给出解存在的结论.

注 2. 条件 1)—4)并不能保证球 $\|x-x_1\|\leqslant\bar{r}_0$ 包含在球 Ω_0 中. 因此就应当假定球 Ω_0 的半径大到能将球 $\|x-x_1\|\leqslant\bar{r}_0$ 包含在 Ω_0 中的程度. 例如，应假定
$$r\geqslant\eta+\bar{r}_0=\eta+\frac{1-\sqrt{1-2h_1}}{h_1}\,\frac{\eta_1}{1-K\eta}.$$

2.5. 如果在定理 1.6 注 1 的条件中，算子 $[P'(x)]^{-1}$ 的范数估计不仅在 x_0 处知道而且在整个区域 Ω_0 上都知道，则加于 h 的限定条件可以减弱，即不需要 $h\leqslant\frac{1}{2}$ 而只要 $h<2$ 就行了(见 Мысовских[1]).

定理 5(Мысовских). 设条件

1) $\|P(x_0)\|\leqslant\eta$;

2) 对于 $x\in\Omega_0$ 存在连续线性算子 $\Gamma(x)=[P'(x)]^{-1}$，并且
$$\|\Gamma(x)\|\leqslant B\qquad(x\in\Omega_0);$$

3) $\|P''(x)\|\leqslant K\qquad(x\in\Omega_0)$

满足. 那么，如果
$$h=B^2K\eta<2\tag{8}$$

并且

$$r > r' = B\eta\sum_{k=0}^{\infty}\left(\frac{h}{2}\right)^{2^k-1},$$

则方程(1)有解 $x^*\in\Omega_0$，而且以 x_0 为初值的基本 Newton 过程收敛于 x^*. 这时，收敛速度由下式给出:

$$\|x^*-x_n\|\leqslant B\eta\frac{\left(\frac{h}{2}\right)^{2^n-1}}{1-\left(\frac{h}{2}\right)^{2^n}}$$

$$(x_{n+1}=x_n-\Gamma(x_n)(P(x_n));\quad n=0,1,\cdots). \tag{9}$$

证. 首先来说明 Newton 过程是可实现的，亦即 $x_n\in\Omega_0$ $(n=0,1,\cdots)$. 事实上，

$$\|x_1-x_0\|=\|-\Gamma(x_0)(P(x_0))\|\leqslant B\eta, \tag{10}$$

$$\|P(x_1)\|=\|P(x_0)+P'(x_0)(x_1-x_0)+\int_{x_0}^{x_1}P''(x)(x_1-x,\cdot)dx\|$$

$$\leqslant\int_0^{B\eta}K(B\eta-t)\,dt=\frac{KB^2\eta^2}{2}=\frac{h\eta}{2}=\eta_1. \tag{11}$$

由(10)推知 $x_1\in\Omega_0$.

我们来计算对应于点 x_1 的 h 值:

$$h_1=B^2K\eta_1=\frac{B^2K\eta h}{2}=\frac{h^2}{2}. \tag{12}$$

设已经证得 $x_n\in\Omega_0$. 以 h_k 和 $\eta_k(k=0,1,\cdots)$ 表示对应于点 x_k 的 h 值和 η 值，则有类似于(10)，(11)与(12)的下列表示式:

$$\|x_{n+1}-x_n\|\leqslant B\eta_n,$$

$$\eta_{k+1}=\frac{h_k\eta_k}{2}\qquad(k=0,1,\cdots,n-1), \tag{13}$$

$$h_k=2\left(\frac{h_{k-1}}{2}\right)^2=\cdots=2\left(\frac{h_0}{2}\right)^{2^k}=2\left(\frac{h}{2}\right)^{2^k}\qquad(k=0,1,\cdots,n).$$

由后两个关系式求出

$$\eta_n = \frac{h_{n-1}\eta_{n-1}}{2} = \frac{h_{n-1}h_{n-2}\eta_{n-2}}{2^2} = \cdots = \frac{h_{n-1}h_{n-2}\cdots h_0}{2^n}\eta_0$$

$$= \left(\frac{h}{2}\right)^{2^n-1}\eta.$$

将此结果代入(13)式，得到

$$\|x_{n+1} - x_n\| \leqslant B\eta\left(\frac{h}{2}\right)^{2^n-1}. \tag{14}$$

由此，利用对$\|x_{k+1} - x_k\|$ $(k=0, 1, \cdots, n-1)$的类似的估计，有

$$\|x_{n+1} - x_0\| \leqslant \sum_{k=0}^{n}\|x_{k+1} - x_k\| \leqslant B\eta\sum_{k=0}^{n}\left(\frac{h}{2}\right)^{2^k-1} \leqslant r' \leqslant r,$$

因而可知 $x_{n+1} \in \Omega_0$.

同样，根据(14)，有

$$\|x_{n+p} - x_n\| \leqslant \sum_{k=n}^{n+p-1}\|x_{k+1} - x_k\| \leqslant \sum_{k=n}^{n+p-1} B\eta\left(\frac{h}{2}\right)^{2^k-1} \xrightarrow[n\to\infty]{} 0, \tag{15}$$

因而序列$\{x_n\}$自收敛，这表明存在 $x^* = \lim\limits_{n\to\infty} x_n$. 易见 x^* 是方程(1)的解.

在(15)中令 $p \to \infty$，得到

$$\|x^* - x_n\| \leqslant B\eta\sum_{k=n}^{\infty}\left(\frac{h}{2}\right)^{2^k-1} \leqslant B\eta\left(\frac{h}{2}\right)^{2^n-1}\sum_{k=0}^{\infty}\left(\frac{h}{2}\right)^{2^n k}$$

$$= B\eta\frac{\left(\frac{h}{2}\right)^{2^n-1}}{1-\left(\frac{h}{2}\right)^{2^n}},$$

这就给出了收敛速度的估计(9).

2.6. 常常有可能将已给出的方程(1)用接近于它的但却简单得多的方程(一般说来仍是非线性的)来代替. 我们来看看在什么

样的条件下根据近似方程的解可以判断方程(1)的可解性.

设方程(1)具有如下形式

$$P(x) \equiv \pi(x) + R(x) = \mathbf{0}. \tag{16}$$

假定元素 x_0 是简化方程的解, 即

$$\pi(x_0) = \mathbf{0}.$$

如果下述条件满足:

1) $\|[\pi'(x_0)]^{-1}(R(x_0))\| \leqslant \eta$;

2) $\|[\pi'(x_0)]^{-1}R'(x_0)\| \leqslant \alpha$

3) $\|[\pi'(x_0)]^{-1}\pi''(x)\| \leqslant K, \|[\pi'(x_0)]^{-1}R''(x)\| \leqslant L (x \in \Omega_0)$.

此外, 如果

$$h = \frac{\eta(K+L)}{(1-\alpha)^2} \leqslant \frac{1}{2} \tag{17}$$

并且

$$r \geqslant r_0 = \frac{1-\sqrt{1-2h}}{h} \frac{\eta}{1-\alpha}, \tag{18}$$

则方程(16)有解 $x^* \in \Omega_0$.

利用定理 1, 在其中令 $\Gamma = [\pi'(x_0)]^{-1}$ 则立即可得到这一结论. 同样, 借助于定理 1 可以建立解 x^* 的唯一性区域.

上面我们只考察了一种特殊的情况. 更一般的情况是方程(1)的右端依赖于若干个参数而且对于其中一个参数值方程的解是已知的, 需要对趋近于初值的参数确定解的存在性. 我们只限于参数是线性地出现在方程中的情形. 确切地讲, 我们将考察形如

$$P(x, \mu) \equiv \pi(x) + \mu R(x) = \mathbf{0} \tag{19}$$

的方程, 这里的 μ 是空间 $\mathbf{Y}$ 中的线性算子, 即 $\mu \in B(\mathbf{Y}, \mathbf{Y})$. 特别, μ 可以是数值因子.

定理 6. 设如下条件满足:

1) $P(x_0, \mathbf{0}) = \pi(x_0) = \mathbf{0}$;

2) 存在连续线性算子 $\Gamma_0=[\pi'(x_0)]^{-1}$ 并且 $\|\Gamma_0\|\leqslant B$;

3) $\|R(x_0)\|\leqslant\eta$, $\|R'(x_0)\|\leqslant\alpha$;

4) $\|\pi''(x)\|\leqslant K$, $\|R''(x)\|\leqslant L\quad(x\in\Omega_0)$.

此外，如果对参数 μ 成立如下关系式

$$h_\mu=\frac{B^2\eta(K+L\|\mu\|)\|\mu\|}{(1+\alpha B\|\mu\|)^2}\leqslant\frac{1}{2},\quad \alpha B\|\mu\|<1, \tag{20}$$

而且球 Ω_0 的半径充分大，则方程(19)有解 $x^*\in\Omega_0$.

证. 这个定理的证明由于方程(16)的结论可直接从以上所指出的假定得出来.

解 $x^*(\mu)$ 的分布区域及其唯一性区域可借助于定理 1 来求. 从定理 1 还推知：以 x_0 为初值的基本Newton过程和简化Newton过程都收敛于解 $x^*(\mu)$. 此外，根据定理 2，序列 $\{\tilde{x}_n(\mu)\}$:

$$\tilde{x}_{n+1}(\mu)=\tilde{x}_n(\mu)-[\pi'(x_0)]^{-1}(P(\tilde{x}_n(\mu),\mu))$$
$$(n=0,1,\cdots;\ \tilde{x}_0(\mu)=x_0) \tag{21}$$

也收敛于 $x^*(\mu)$. 过程(21)具有这样的优点，它用了对应于初值点 x_0 的逆算子及初始参数值 $\mu=\mathbf{0}$.

注 1. 应注意的是，对每个 μ 来说，实方程

$$\left(\frac{K}{2}t^2-\frac{1}{B}t\right)+\|\mu\|\left(\frac{L}{2}t^2+\alpha t+\eta\right)=0$$

都是(19)的控制方程，而(20)不是别的，正是这个方程有实根的条件.

注 2. 某些更复杂的讨论使得我们能够对一般情形即对参数 μ 是非线性地出现在方程 $P(x,\mu)=\mathbf{0}$ 中的情形加以分析.

可以指出，如果不是对 x_0 和 Γ_0 而是对 x 和 Γ 应用定理 1，则在对参数值的更宽的限制之下可以得到方程组

$$P(x,\mu)=\mathbf{0},\quad \Gamma P'(x,\mu)=I$$

解的精确到参数 μ 的“一次幂”的表达式.

我们只限于参数取数值而近似方程 $\pi(x)=0$ 是线性方程的情形加以讨论．也就是说，我们只讨论方程(19)的特殊情况，即讨论形如

$$P(x,\mu)\equiv U(x-x_0)+\mu R(x)=0 \tag{22}$$

的方程，其中 $U=(\pi'(x_0))$ 是连续线性算子．

这时，所指出的解(精确到参数 μ 的一次幂)为

$$x_1=x_0-\mu\Gamma_0(R(x_0)),\quad \Gamma_1=\Gamma_0-\mu\Gamma_0R'(x_0)\Gamma_0.$$

因此就需要估计三个值：

1) $\|\Gamma_1(P(x_1,\mu))\|=\|[\Gamma_0-\mu\Gamma_0R'(x_0)\Gamma_0][-\mu R(x_0)+\mu R(x_0-\mu\Gamma_0R(x_0))]\|\leqslant\eta_1;$

2) $\|\Gamma_1P'(x_1,\mu)-I\|=\|[\Gamma_0-\mu\Gamma_0R'(x_0)\Gamma_0][U+\mu R'(x_0-\mu\Gamma_0R(x_0))]-I\|\leqslant\delta_1;$

3) $\|[\Gamma_0-\mu\Gamma_0R'(x_0)\Gamma_0]R''(x)\|\leqslant L_1 \qquad (x\in\Omega_0).$

由此可知方程(22)当

$$\frac{L_1\eta_1|\mu|}{(1-\delta_1)^2}\leqslant\frac{1}{2}$$

时可解．

2.7. 在参数变化的一定的区域内解的存在性和求解过程的收敛性使得我们还有可能作出关于解 $x^*(\mu)$ 对参数的依赖性的特征的结论．例如，由于 $\tilde{x}_0(\mu)=x_0$ 连续地依赖于 μ 而 $\pi(x)$ 和 $R(x)$ 又是 x 的连续函数，所以(21)中所有 $\tilde{x}_n(\mu)$ 都是参数的连续函数，也就是说 $x(\mu)$ 是连续地依赖于 μ 的．

在一定的条件下，还可以判断解 $x^*(\mu)$ 对参数依赖性的解析特征．

在空间 $\mathbf{X}$ 中取值的函数 $x(\mu)$ 称作是弱解析的，如果对任何线性泛函 $f\in\mathbf{X}^*$ 函数 $f(x(\mu))$ 在通常意义下是解析的．

算子 P 称作是解析的，如果它将任何弱解析函数 $x(\mu)$ 同样变

成空间 Y 中的弱解析函数 $P(x(\mu))$.

设方程(19)中的 μ 是数值参数并假定 π 和 R 是解析算子. 那么, 容易验证(21)所有近似解都是 μ 的解析函数. 设 f 是空间 X 中任意的连续线性泛函. 由于元素 $\tilde{x}_n(\mu)$ 的范数关于 μ 的一致有界性, 所以解析函数族

$$\varphi_n(\mu)=f(\tilde{x}_n(\mu)) \qquad (n=0,1,\cdots)$$

是总体有界的, 因而函数

$$\varphi(\mu)=\lim_{n\to\infty}\varphi_n(\mu)=f(x^*(\mu))$$

是解析的. 由于 f 是任意的泛函, 所以 $x^*(\mu)$ 是 μ 的弱解析函数.

如果以自然的方式将函数 $x^*(\mu)$ 定义为关于 μ 的幂级数

$$x^*(\mu)=x_0+\mu x_1+\cdots+\mu^n x_n+\cdots,$$

则由以上所指出的可推知这个级数对 $|\mu|<\mu_0$ 的所有的 μ 应该是弱收敛的.

Newton 法的各种应用和推广已有大量的著作, 特别是 Красносельский 和 Коллатц 等人的著作.

§3. Newton法对具体泛函方程的应用

3.1. 先来考察 Newton 法对一个实的或复的方程

$$f(z)=0 \tag{1}$$

的应用.

利用定理 1.6 并考虑到定理中所出现的量 η 和 K 具有如下意义:

$$\eta\geqslant\frac{|f(z_0)|}{|f'(z_0)|}, \qquad K\geqslant\max\frac{|f''(z)|}{|f'(z_0)|}$$

(其中 z_0 是初次逼近), 如果

$$h=\eta K\leqslant\frac{1}{2} \quad 或者 \quad \frac{|f(z_0)||f''(z_0)|}{|f'(z_0)|^2}\leqslant\frac{1}{2},$$

则得知(1)有根，这个根位于圆

$$|z-z_0|\leqslant r_0=\frac{1-\sqrt{1-2h}}{h}\eta \tag{2}$$

中，并且根在圆

$$|z-z_0|<r_1=\frac{1+\sqrt{1-2h}}{h}\eta^{*)} \tag{3}$$

中是唯一的.

Newton 法还可用于求解代数方程组

$$\varphi_j(\xi_1, \xi_2, \cdots, \xi_m)=0 \qquad (j=1, 2, \cdots, m), \tag{4}$$

方程组(4)可看成m维空间 $\mathbf{X}$ 中的一个方程

$$P(x)=\mathbf{0}$$

即 $y=P(x)$，而

$$y=(\varphi_1(\xi_1, \cdots, \xi_m), \cdots, \varphi_m(\xi_1, \cdots, \xi_m)),$$
$$x=(\xi_1, \cdots, \xi_m).$$

设 $x_0=(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_m^{(0)})$是初始逼近，将它代入 Newton 法的方程

$$P'(x_0)(x-x_0)+P(x_0)=\mathbf{0}$$

中并考虑到在 XVII. 1. 6 中所指出的导数 $P'(x_0)$的表达式，于是我们就得到关于修正量 $\Delta x=x_1-x_0=(\Delta\xi_1, \Delta\xi_2, \cdots, \Delta\xi_m)$的线性方程组

$$\sum_{k=1}^{m}\frac{\partial\varphi_j(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_m^{(0)})}{\partial\xi_k}\Delta\xi_k$$
$$=-\varphi_j(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_m^{(0)})$$
$$(j=1, 2, \cdots, m). \tag{5}$$

由这个方程组求出 Δx，因而也就求得 x_1. 按同样的办法求 x_2,

---

*) 这时假定二阶导数是在圆 $|z-z_0|\leqslant r$ 中估计的，其中，对于方程(2)来说 $r\geqslant r_0$，对于式(3)来说 $r\geqslant r_1$.

等等.

如果使用的是简化的 Newton 过程,则方程(5)左端的矩阵在每一步都不变,所变的只是等式的右端.

Newton 法的收敛性条件依赖于赋范空间 $\mathbf{X}$ 如何选取. 我们考察两种情况: $\mathbf{X}=\mathbf{l}_m^\infty$ 和 $\mathbf{X}=\mathbf{l}_m^2$.

定理 1. 设函数 $\varphi_1, \varphi_2, \cdots, \varphi_m$ 和初始逼近 x_0 满足以下条件:

1) $|\varphi_j(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_m^{(0)})| \leqslant \eta' \qquad (j=1, 2, \cdots, m);$

2) Jocobi 矩阵

$$\left(\frac{\partial \varphi_j}{\partial \xi_k}(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_m^{(0)})\right) \qquad (j, k=1, 2, \cdots, m)$$

的行列式 $D\neq 0$, 而且, 如果用 $A_{j,k}$ 表示元素 $\dfrac{\partial \varphi_j}{\partial \xi_k}$ 的代数余子式, 则还成立如下估计式

$$\frac{1}{|D|}\sum_{j=1}^{m}|A_{j,k}| \leqslant B' \quad (k=1, 2, \cdots, m);$$

3) $\left|\dfrac{\partial^2 \varphi_j}{\partial \xi_k \partial \xi_s}(\xi_1, \xi_2, \cdots, \xi_m)\right| \leqslant L$

$$(j, k, s=1, 2, \cdots, m;\ |\xi_i - \xi_i^{(0)}| \leqslant r);$$

4) $h = B'^2\eta' L m^2 \leqslant \dfrac{1}{2}.$

此外, 如果

$$r \geqslant \frac{1-\sqrt{1-2h}}{h}B'h' \tag{6}$$

则方程组(4)在 x_0 点的邻域内存在解 $x^*=(\xi_1^*, \xi_2^*, \cdots, \xi_m^*)$ 而且基本 Newton 过程和简化 Newton 过程均收敛于该解.

定理的证明归结为对定理 1.6 诸条件的验证, 确切地讲, 如果取 $\mathbf{X}=\mathbf{l}_m^\infty$, 就归结为对该定理的注 1 的条件的验证. 我们建议读

者去作这种验证*).

注 1. 基本 Newton 过程和简化 Newton 过程的收敛速度由定理 1.6 的不等式(33)和(34)可给出估计.

注 2. 因为当$0<h\leqslant\frac{1}{2}$时

$$\frac{1-\sqrt{1-2h}}{h}\leqslant 2,$$

所以，若取 $r=2B'\eta'$，则(6)满足.

注 3. 设方程组(4)的方程个数 $m=2$. 那么

$$A_{j,k}=\pm\frac{\partial\varphi_{3-k}}{\partial\xi_{3-j}}(\xi_1^{(0)},\xi_2^{(0)}) \qquad (j,k=1,2),$$

因此，如果已知估计

$$\left|\frac{\partial\varphi_j}{\partial\xi_k}(\xi_1^{(0)},\xi_2^{(0)})\right|\leqslant l \qquad (j,k=1,2),$$

则作为 B' 可取为

$$B'=\frac{2l}{|D|},$$

这时条件 4)可写为

$$\frac{16l^2\eta' L}{|D|}\leqslant\frac{1}{2}$$

或者

$$32l^2\eta' L\leqslant|D|.$$

这个条件是由 A. Ostrowski 得到的. 有趣的是，要是直接证明 Newton 法的收敛性，即使这种最简单的情形也是十分复杂的.

如果取赋范空间 **X** 为 $\mathbf{l}_m^2$，则定理 1 的条件 1)—4)成为

$$1')\quad \|\Gamma_0(P(x_0))\|=\|\Delta x\|=\left[\sum_{k=1}^{m}|\Delta\xi_k|^2\right]^{\frac{1}{2}}\leqslant\eta;$$

---

*) 容易证明 $\|P''(x)\|\leqslant Lm^2$ (见 XVII. 2.2).

2′) $\|\Gamma_0\|=\sqrt{\Lambda}\leqslant B'$;

这里 Γ_0 象以前那样总表示$[P'(x_0)]^{-1}$, Λ 表示矩阵 $\Gamma_0\Gamma_0^*$ 的最大特征值;作为 B' 可取为,例如

$$B'=\frac{1}{|D|}\left[\sum_{j,k=1}^{m}|A_{j,k}|^2\right]^{\frac{1}{2}};$$

3′) $\|P''(x)\|\leqslant\left[\sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{s=1}^{m}\left|\frac{\partial\varphi_j}{\partial\xi_k\partial\xi_s}\right|^2\right]^{\frac{1}{2}}\leqslant K'$,

因而作为K'可取为:

$$K'=Lm\sqrt{m};$$

4′) $h=B'K'\eta\leqslant\frac{1}{2}$.

这些条件的满足就保证了方程组(4)的解的存在以及基本 Newton 过程和简化 Newton 过程收敛于(4)的解.

对于以上所述,我们以一个例子来说明. 考察方程组

$$3\xi_1^2\xi_2+\xi_2^2=1,$$
$$\xi_1^4+\xi_1\xi_2^3=1.$$

初始逼近 $x_0=(\xi_1^{(0)},\xi_2^{(0)})$ 选取为

$$\xi_1^{(0)}=0.98,\ \xi_2^{(0)}=0.32.$$

确定修正的方程组是

$$1.880\Delta\xi_1+3.188\Delta\xi_2=0.045,$$
$$3.798\Delta\xi_1+0.301\Delta\xi_2=0.046,$$

由此我们求得

$$\Delta\xi_1=0.0105,\quad \Delta\xi_2=0.0075.$$

由此得到

$$-\Gamma_0=\begin{pmatrix}-0.026 & 0.256 \\ \end{pmatrix}.$$

$$\|\Gamma_0\| = 0.492 < 0.5 = B',$$

$$\|P(x_0)\| = \max[0.046, 0.045] < 0.05 = \eta',$$

现对$\|x-x_0\| \leqslant 2B'\eta'$ 也就是对 $0.93 \leqslant \xi_1 \leqslant 1.03$, $0.27 \leqslant \xi_2 \leqslant 0.37$ 来估计函数 φ_1 和 φ_2 的二阶导数. 因为

$$\frac{\partial^2\varphi_1}{\partial\xi_1^2} = 6\xi_2, \quad \frac{\partial^2\varphi_1}{\partial\xi_1\partial\xi_2} = 6\xi_1, \quad \frac{\partial^2\varphi_1}{\partial\xi_2^2} = 6\xi_2,$$

$$\frac{\partial^2\varphi_2}{\partial\xi_1^2} = 12\xi_1^2, \quad \frac{\partial^2\varphi_2}{\partial\xi_1\partial\xi_2} = 3\xi_2^2, \quad \frac{\partial^2\varphi_2}{\partial\xi_2^2} = 6\xi_1\xi_2,$$

所以作为 L 可以取为:

$$L = 12.8 > \max\frac{\partial^2\varphi_2}{\partial\xi_1^2} = 12(1.03)^2,$$

当 B', η' 和 L 取上述值时, h 的值将为

$$h = B'^2\eta' L m^2 = 0.25 \cdot 0.05 \cdot 12.8 \cdot 4 = 0.64.$$

因此定理 1 并不能给出关于 Newton 法收敛性的结论. 这是因为对 $\|P''(x)\|$ 的估计太粗糙了. 事实上, 我们取

$$\|P''(x)\| \leqslant Lm^2 = 51.2,$$

其实存在着更精细的估计

$$\|P''(x)\| \leqslant 16 = K'.$$

利用这一估计得到

$$h = B'^2\eta' K' = 0.25 \cdot 0.05 \cdot 16 = 0.2,$$

这就可以作出 Newton 过程收敛性的肯定结论了.

如果直接利用定理 1.6, 则对 h 还可得到更小的值. 事实上,

$$\|\Gamma_0(P(x_0))\| = \|\Delta x\| = \max[|\Delta\xi_1|, |\Delta\xi_2|] = 0.0105,$$

而

$$\|\Gamma_0 P''(x)\| < 7.2.$$

因此, 取 $\eta = 0.0105$, $K = 7.2$, 我们就得到

$$h = \eta K < 0.08. \tag{7}$$

利用简化的 Newton 过程, 可得到逐次逼近

$$\xi_1^{(1)} = 0.9905, \qquad \xi_2^{(1)} = 0.3275,$$

$$\xi_1^{(2)} = 0.99117, \qquad \xi_2^{(2)} = 0.32738,$$

$$\xi_1^{(3)} = 0.991189, \qquad \xi_2^{(3)} = 0.327382;$$

按定理 1.6 的公式 (34) 所算得的第三次逼近的误差估计得到的值为 0.0008

(这时 h 的值是按(7)式取的).

如果应用定理 2.4, 则还可以指出更精确的误差估计; 亦即

$$0.991173 \leqslant \xi_1^* \leqslant 0.991205, \qquad 0.327366 \leqslant \xi_2^* \leqslant 0.327398.$$

因此第三逼近精确到小数后 4 位.

如果利用赋范空间 $\mathbf{l}_2^2$, 则从条件 1′)—3′)可得到

$$\|\Gamma_0\| \leqslant 0.448 < 0.5 = B',$$
$$\|\Gamma_0(P(x_0))\| = \|\Delta x\| = 0.0133 < 0.015 = \eta,$$
$$\|P''(x)\| < 15.2 = K', \quad h = B'K'\eta = 0.1026.$$

3.2. 我们现在来考察 Newton 法对非线性积分方程的应用. 设给定积分方程

$$x(s) = \int_0^1 K(s, t, x(t))\,dt, \tag{8}$$

其中 $K(s, t, u)$是s, t 和 u 的连续函数而且具有所需要的各阶的连续导数. 在泛函空间 $\mathbf{X}$ 中引入算子 P:

$$y = P(x), \; y(s) = x(s) - \int_0^1 K(s, t, x(t))\,dt \tag{9}$$

并将方程(8)写为泛函方程

$$P(x) = \mathbf{0}$$

的形式. 对于这个方程的 Newton 过程按如下方式来构造: 设 x_0 是初始逼近; 我们假定空间 $\mathbf{X}$ 和函数$K(s, t, u)$的组成和性质能保证 $P'(x_0)$ 用"在积分号下求导数"的方法得到, 也就是说, $z = P'(x_0)(x)$表示

$$z(s) = x(s) - \int_0^1 K'_u(s, t, x_0(t))x(t)\,dt; \tag{10}$$

修正量 $\Delta x = x_1 - x_0$ 由方程

$$P'(x_0)(\Delta x) = -P(x_0)$$

确定. 由于(10), 这个方程具有如下形式

$$\Delta x(s) - \int_0^1 K'_u(s, t, x_0(t))\Delta x(t)\,dt = \varepsilon_0(s), \tag{11}$$

其中

$$\varepsilon_0(s)=\int_0^1 K(s,t,x_0(t))dt-x_0(s)$$

是方程(8)与初始逼近相对应的误(残)差(*Невязка*).

因此,为了求每一个下步逼近就需要解线性积分方程. 这时,如果利用简化的 Newton 过程,则核函数在每一步中都不改变.

对于牛顿过程的收敛性,根据一般定理可以指出种种不同的结果,这要看基本空间是选取什么样的赋范空间了. 于是,如果取 $\mathbf{X}=\mathbf{C}[0,1]$,我们就得到这样的定理.

定理 2. 设如下条件满足:

1) 积分方程(11)对于核函数 $K(s,t)=K'_u(s,t,x_0(t))$ 存在着预解式 $G(s,t)$,并且

$$\int_0^1 |G(s,t)|dt\leqslant B \qquad (s\in[0,1]); \tag{12}$$

2) $|\varepsilon_0(s)|\leqslant\eta' \qquad (s\in[0,1])$;

3) $\int_0^1 |K''_{u^2}(s,t,u)|dt\leqslant K'$ 在由不等式

$$0\leqslant s\leqslant 1, \quad |u-x_0(s)|\leqslant 2(1+B)\eta'$$

所定义的 s 和 u 的区域中成立;

4) $h=(1+B)^2K'\eta'\leqslant\frac{1}{2}$.

则基本 Newton 过程和简化 Newton 过程均收敛于方程(8)的解 $x^*(s)$. 解的存在域是

$$|x^*(s)-x_0(s)|\leqslant\frac{1-\sqrt{1-2h}}{h}(1+B)\eta' \qquad (s\in[0,1]),$$

解的唯一性区域是

$$0\leqslant s\leqslant 1, \quad |u-x_0(s)|\leqslant\frac{1+\sqrt{1-2h}}{h}(1+B)\eta'.$$

如果注意到由于定理 XVII.3.2，算子P的导数$P'(x_0)$具有(10)的形式，而且$P''(x)$是以$K''_{u^2}(s,t,x_0(t))$为核函数的双线性积分算子的话，则不难验证定理1.6的所有条件均满足. 至于算子$\Gamma_0=[P'(x_0)]^{-1}$，则正如在 XIII.6.1 中所指出的那样，具有如下形式

$$z=\Gamma_0(y),\quad z(s)=y(s)+\int_0^1 G(s,t)y(t)dt,$$

因而

$$\|\Gamma_0\|\leqslant 1+B.$$

如果取$\mathbf{L}^2(0,1)$为$\mathbf{X}$，则依据定理 XIII.3.1 还可指出积分方程(8)可解性的其他条件. 也就是

1') $\int_0^1|\varepsilon_0(s)|^2ds\leqslant\eta'^2$；

2') 不等式

$$\left|\frac{\lambda_n}{1-\lambda_n}\right|\leqslant B'\qquad(n=1,2,\cdots)$$

成立，其中λ_n是当核函数$K(s,t)=K'_u(s,t,x_0(t))$对称时这个核函数的特征值(如果核$K(s,t)$不对称，要求就复杂一些，见 V.2.6 和 V.2.7)；

3') $|K''_{u^2}(s,t,u)|\leqslant K'\qquad(s,t\in[0,1],-\infty<u<\infty)$；

4') $h=B'^2K'\eta'\leqslant\frac{1}{2}$.

3.3. 定理2的条件或条件1')—4')都假定存在着与精确解充分接近的初始逼近x_0，我们现在来考察构造这种初始逼近的相当一般的方法，这个方法的任务就在于用一个结构更简单的而且它的求解归结为求解代数方程组的方程去代替已知的积分方程(8).

亦即，我们假定核$K(s,t,u)$可用更简单的形如

$$H(s,t,u)=\sum_{k=1}^{m}h_k(t,u)\omega_k(s)$$

的核来逼近，其中 $\omega_k(s)(k=1,2,\cdots,m)$ 是某些函数，不失一般性，不妨认为是两两正交规格化的函数系．例如，可取 H 为函数 K 的按完全正交函数系 $\{\omega_k(s)\}(k=1,2,\cdots)$ 的 Fourier 级数展开式的部分和．

方程

$$x(s)=\int_0^1 H(s,t,x(t))\,dt \tag{13}$$

自然看作对于方程(8)的逼近方程．但方程(13)的解 $x_0(t)$ 具有如下形式：

$$x_0(t)=\sum_{k=1}^{m}A_k\omega_k(t),$$

其中系数 $A_k(k=1,2,\cdots,m)$ 由非线性代数方程组

$$A_j=\int_0^1 h_j\left(t,\sum_{k=1}^{m}A_k\omega_k(t)\right)dt \qquad (j=1,2,\cdots,m) \tag{14}$$

确定．我们正是将 $x_0(t)$ 取作方程(8)的初始逼近．

因为在很多情形下重要的只是关于方程(8)的解的存在，解的分布区域以及解的唯一性问题，所以就很自然地应用到定理 2.6 这一格式．为此取 $\mathbf{X}=\mathbf{C}[0,1]$ 并将方程(8)写成

$$P(x)\equiv\pi(x)+\mu R(x)=\mathbf{0},$$

$$\pi(x)(s)=x(s)-\int_0^1 H(s,t,x(t))\,dt,$$

$$R(x)(s)=\int_0^1[K(s,t,x(t))-H(s,t,x(t))]\,dt,$$

$$\mu=1.$$

在这些记法之下，可指出以下定理．

定理3. 设如下条件满足:

1) 存在关于核函数 $H'_u(s,t,x_0(t))$ 的像解式 $G_0(s,t)$, 并且

$$\int_0^1 |G_0(s,t)|\,dt \leqslant B \qquad (s\in[0,1]);$$

2) $$\left|\int_0^1 [K(s,t,x_0(t))-H(s,t,x_0(t))]\,dt\right| = |\varepsilon_0(s)| \leqslant \eta \qquad (s\in[0,1]);$$

3) $$\int_0^1 |K'_u(s,t,x_0(t))-H'_u(s,t,x_0(t))|\,dt \leqslant \alpha \qquad (s\in[0,1]);$$

4) $$\int_0^1 |H''_{u^2}(s,t,u)|\,dt \leqslant K \qquad (s\in[0,1],\ |u-x_0(s)|\leqslant r);$$

5) $$\int_0^1 |K''_{u^2}(s,t,u)-H''_{u^2}(s,t,u)|\,dt \leqslant L \qquad (s\in[0,1],\ |u-x_0(s)|\leqslant r);$$

6) $$h=\frac{(1+B)^2(K+L)\eta}{[1-\alpha(B+1)]^2}\leqslant\frac{1}{2},\quad \alpha(B+1)<1.$$

这时, 如果

$$r\geqslant r_0=\frac{1-\sqrt{1-2h}}{h}\cdot\frac{\eta(1+B)}{1-\alpha(1+B)},$$

则方程(8)有解 x^*, 并且

$$|x^*(s)-x_0(s)|\leqslant r_0. \tag{15}$$

与定理2相比, 这个定理的方便之处就在于核 $H'_u(s,t,x_0(t))$ 是退化的, 因而像解式 $G_0(s,t)$ 容易求得.

为了确定方程(8)的解, 可利用2.6中所指出的过程, 但是这时收敛性要比利用简化的 Newton 过程慢一些.

为了说明以上所讲的, 我们来考察一个例子. 设给定了方程

$$x(s)=\frac{1}{2}\int_0^1 [x(t)]^2 \sin st\,dt+1.$$

作为逼近核可取

$$H(s,t,u)=\frac{1}{2}stu^2+1.$$

因此这时方程(13)应是

$$x(s)=\frac{1}{2}\int_0^1[x(t)]^2stdt+1.$$

这个方程的解是 $x_0(t)=1+At$ 而且 $A=0.405887$ (是由二次方程求得的).不难算出定理 3 中所出现的常数分别为

$$B=1.124,\quad \eta=0.0367,\quad \alpha=0.055,\quad K=0.5,\quad L=0.0417.$$

因此

$$0.1<h<0.103;\quad \alpha(1+B)<0.12,$$

此处,由(15)得

$$|x^*(s)-x_0(s)|<0.096\qquad (s\in[0,1]).$$

如果我们取

$$H(s,t,u)=\frac{1}{2}\left(st-\frac{s^3t^3}{6}\right)u^2+1,$$

则会得到

$$x_0(s)=1+0.38617s-0.0345s^3,$$

而估计式(15)将变为

$$|x^*(s)-x_0(s)|<0.0119\qquad (s\in[0,1]).$$

3.4. 现在来看看 Newton 法对微分方程的应用. 设给定微分方程

$$x'(t)-\varphi(x(t),t)=0\qquad (x(0)=0).\tag{16}$$

设 $\mathbf{C}^1$ 是由在区间$[0,a]$上连续的而在 $t=0$ 时等于零的所有函数构成的空间,我们在这个空间中用

$$\|x\|=\max_{t\in[0,a]}|x(t)|+\lambda\max_{t\in[0,a]}|x'(t)|$$

引进范数,其中 λ 是一个我们以后再来确定的正系数. 假定函数 $\varphi(u,t)$在区域

$$0\leqslant t\leqslant a;\qquad |u-x_0(t)|\leqslant\delta\qquad (x_0\in\mathbf{C}^1)$$

上连续并且有关于 u 的连续的二阶导数,我们来考察算子 P:

$$y=P(x),\quad y(t)=x'(t)-\varphi(x(t),t).$$

不难验证P映球$\|x-x_0\|\leqslant\delta$(用Ω表示这个球)到空间$C[0, 1]$并且在Ω中有连续的一阶和二阶导数. 在这样的假定下就有

$$P'(z)(x)(t)=x'(t)-\varphi_u'(z(t), t)x(t), \tag{17}$$

$$P''(z)(x, \tilde{x})(t)=-\varphi_{u^2}''(z(t), t)x(t)\tilde{x}(t). \tag{18}$$

我们来求$\Gamma_0=[P'(x_0)]^{-1}$. 由于(17), 元素$x=\Gamma_0(y)$是微分方程

$$x'(t)=\varphi_u'(x_0(t), t)x(t)+y(t) \tag{19}$$

的解, 也就是说

$$x(t)=\psi(t)\int_0^t\frac{y(s)}{\psi(s)}ds,$$

其中

$$\psi(t)=\exp\left(\int_0^t\varphi_u'(x_0(s), s)ds\right).$$

因此,

$$\max_{t\in[0,a]}|x(t)|\leqslant a\frac{\max|\psi(t)|}{\min|\psi(t)|}\|y\|,$$

而从方程(19)得到

$$\max_{t\in[0,a]}|x'(t)|\leqslant\max_{t\in[0,a]}|\varphi_u'(x_0(t), t)|\max_{t\in[0,a]}|x(t)|+\theta\|y\|$$
$$\leqslant\theta\|y\|.$$

这样, 就有

$$\|x\|\leqslant\left[a\frac{\max|\psi(t)|}{\min|\psi(t)|}+\lambda\theta\right]\|y\|,$$

因此

$$\|\Gamma_0\|\leqslant a\frac{\max|\psi(t)|}{\min|\psi(t)|}+\lambda\theta.$$

所得到的这个估计使得我们可以利用定理1.6得出如下结果.

定理 4. 设下述条件满足:

1) $|x_0'(t)-\varphi(x_0(t), t)|\leqslant\eta'\qquad(t\in[0, a])$;

2) $|\varphi'_u(x_0(t),t)|\leqslant M_1 \qquad (t\in[0,a])$;

3) $|\varphi''_{u^2}(u,t)\leqslant M_2 \qquad (t\in[0,a],\ |u-x_0(t)|\leqslant\delta)$;

4) $h_0=M_2a^2e^{4aM_1}\eta'<\frac{1}{2}$.

这时，如果

$$\delta>r_0=\frac{1-\sqrt{1-2h_0}}{h_0}e^{2aM_1}a\eta',$$

则方程(16)在区间$[0,a]$中有唯一解 x^*, 并且

$$|x^*(t)-x_0(t)|<r_0.$$

为了验证定理 1.6 的条件, 只要说明

$$\max|\psi(t)|\leqslant e^{aM_1},\quad \min|\psi(t)|\geqslant e^{-aM_1}$$

就够了. 因此该定理中的 B' 可取

$$B'=ae^{2aM_1}+\lambda\theta.$$

此外还可设 $K'=M_2$, 于是, 方程(16)可解性条件可写为

$$h=B'^2K'\eta'=(ae^{2aM_1}+\lambda\theta)^2M_2\eta'\leqslant\frac{1}{2}. \qquad (20)$$

由于 $\lim\limits_{\lambda\to0}h=h_0$, 根据条件 4), 只要取 λ 充分小即可保证(20)成立.

注 1. 对于非零初始条件只须经显然的变量代换即可化成我们这里的情形.

注 2. 这里所用的这个方法还可用于研究微分方程组. 在这种情形使用这个方法的困难之处就在于通常不可能有线性方程组的解析解, 因而不能得到算子 Γ_0 的范数估计.

3.5. 我们现在来考察微分方程

$$x''(t)+x(t)+\mu\varphi(x(t),x'(t),t)=0, \qquad (21)$$

其中 $\varphi(u,v,t)$是连续函数, 存在关于 u 和 v 的直到二阶的连续导数并且是 t 的以 $\omega>0$ 为周期的周期函数.

和 3.4 的讨论类似, 我们指出在什么样的条件下方程 (21)

具有以 ω 为周期的周期解，换言之，也就是说有满足边界条件

$$x(\omega)=x(0),\quad x'(\omega)=x'(0) \tag{22}$$

的解.

这一次我们将利用定理 2.6. 设 $\widetilde{\mathbf{C}}^2$ 是由所有以 ω 为周期的二次连续可微函数所组成的空间. $\widetilde{\mathbf{C}}^2$ 中的范数由等式

$$\|x\|=\max_t|x(t)|+\max_t|x'(t)|+\lambda\max_t|x''(t)|\quad(x\in\widetilde{\mathbf{C}}^2)$$

确定，其中 $\lambda>0$ 将在以后确定. 另外，引进算子 π 和 R:

$$y=\pi(x),\quad y(t)=x''(t)+x(t),$$

$$z=R(x),\quad z(t)=\varphi(x(t),x'(t),t).$$

这时，方程(21)可写为

$$\pi(x)+\mu R(x)=\mathbf{0}. \tag{23}$$

我们来进行为应用定理 2.6 而必要的一些估计.

算子 π 显然是连续的和线性的，因而方程 $\pi(x)=\mathbf{0}$ 的解是元素 $x_0=\mathbf{0}$. 如果 $\cos\omega\neq1$，则存在 $\Gamma_0=\pi^{-1}$. 算子 Γ_0 可通过解微分方程(在条件(22)之下)

$$x''+x=y$$

而求得，我们略去实质上不复杂但令人厌倦的那些计算，直接指出算子 Γ_0 的如下的范数估计:

$$\|\Gamma_0\|\leqslant\omega\left[2+\frac{1+|\cos\omega|+|\sin\omega|}{1-\cos\omega}\right]+\lambda\theta$$

$$\leqslant\frac{4+1+\sqrt{2}}{1-\cos\omega}\,\omega+\lambda\theta<\frac{6.5\omega}{1-\cos\omega}+\lambda\theta;$$

这里的 θ 表示包含 ω 的某个表达式.

因此，B 可取为:

$$B=\frac{6.5\omega}{1-\cos\omega}+\lambda\theta=B_0+\lambda\theta.$$

考虑到

$$R'(x_0)(x)(t)=\varphi'_u(0,0,t)x(t)+\varphi'_v(0,0,t)x'(t),$$

$$\begin{aligned}R''(z)(x,\tilde{x})(t)=&\varphi''_{u^2}(z(t),z'(t),t)x(t)\tilde{x}(t)\\&+\varphi''_{uv}(z(t),z'(t),t)[x(t)\tilde{x}'(t)\\&+x'(t)\tilde{x}(t)]\\&+\varphi''_{v^2}(z(t),z'(t),t)x'(t)\tilde{x}'(t),\end{aligned}$$

现在可以叙述如下的定理.

定理 5. 设如下条件满足:

1) $\omega=2n\pi\qquad(n=1,2,\cdots)$;

2) $|\varphi(0,0,t)|\leqslant\eta$;

3) $|\varphi'_u(0,0,t)|\leqslant\alpha,\quad|\varphi'_v(0,0,t)|\leqslant\alpha$;

4) $|\varphi''_{u^2}(u,v,t)|\leqslant L,\quad|\varphi''_{uv}(u,v,t)|\leqslant L,$

$|\varphi''_{v^2}(u,v,t)|\leqslant L.$

这时,如果

$$\mu<\frac{1-\cos\omega}{6.5\omega\,(\alpha+\sqrt{2L\eta})},\tag{24}$$

则方程(21)具有唯一的以 ω 为周期的周期解.

要证明这个定理只须验证定理 2.6 的不等式(20)当 λ 充分小而且

$$\frac{B_0^2L\eta\mu^2}{(1-\alpha B_0\mu)^2}<\frac{1}{2}\tag{*}$$

时成立. 定理 2.6 的不等式(20)在这里具有下列形式

$$h_\mu=\frac{B^2L\eta\mu^2}{(1-\alpha B\mu)^2}\leqslant\frac{1}{2},$$

而不等式(*)和不等式(24)是等价的.

注. 如果 φ 解析地依赖于其变元, 则根据 2.7 所述可以肯定方程(21)的周期解对满足关系式(24)的 μ 也是解析地相依赖的.

对于更高阶的方程可以类似地讨论.

3.6. 我们来考察 Newton 法对摄动理论的应用.

设U和V是映 Banach 空间 $\mathbf{X}$ 到自身的连续线性算子. 设要求算子 $U_t = U + tV$ 的特征值和相应的特征元，但假定 $t=0$ 时这个问题的解是已知的，也就是说算子U的特征值 λ_0 和对应的特征元 x_0 是已知的.

在对算子U和V作某些补充要求的条件下，我们来证明当t充分小时算子U_t存在分别接近于λ_0和x_0的特征值λ_t和对应的特征元x_t. 亦即，有如下的定理.

定理 6. 设如下条件满足:

1) $\bar{\lambda}_0$ 是算子 U^* 的特征值:

$$U^*(f_0) = \bar{\lambda}_0 f_0,$$

其中 f_0 表示满足规格化条件

$$f_0(x_0) = 1$$

的特征元;

2) 算子 $U-\lambda_0 I$ 看作是从 $\mathbf{X}_0$ 到 $\mathbf{X}$ 的算子并且存在连续逆算子T:

$$T(Ux-\lambda_0 x) = x, \quad UTx - \lambda_0 Tx = x \qquad (x \in \mathbf{X}_0)$$

其中 $\mathbf{X}_0 = \mathbf{N}(f_0)$ 是由所有使得 $f_0(y)=0$ 的 $y \in \mathbf{X}$ 组成的空间. 这时，如果

$$\|T\| \|V\| |t| \leqslant \frac{1}{1+2\|f_0\|\|x_0\| + 2\sqrt{(1+\|f_0\|\|x_0\|)\|f_0\|\|x_0\|}}, \tag{25}$$

则算子U_t存在这样的特征值λ_t，使得当对应的特征元x_t满足规格化条件

$$f_0(x_t) = 1$$

时，就有

$$|\lambda_t - \lambda_0| \leqslant A|t|, \quad \|x_t - x_0\| \leqslant B|t|,$$

其中A和B是由算子U和V所确定的常数.

证. 需要从条件

$$U(x)+tV(x)=\lambda x,\quad f_0(x)=1$$

来确定 λ 和 x. 令

$$x=x_0+z,\quad \lambda=\lambda_0+\delta$$

而引进两个新的未知数 z 和 δ. 它们应该满足条件

$$U(z)+tV(x_0+z)=\delta(x_0+z)+\lambda_0 z,\ f_0(z)=0,$$

也就是说应该在空间 $\mathbf{X}_0$ 中求方程

$$U(z)+tV(x_0+z)=\delta(x_0+z)+\lambda_0 z \tag{26}$$

的解. 为此把泛函 f_0 作用于(26)的两端. 因为

$$f_0(Uz-\lambda_0 z)=U^*(f_0)(z)-(\bar{\lambda}_0 f_0)(z)=0,$$

由此得到用 z 表示 δ 的等式

$$\delta=tf_0(V(x_0+z)), \tag{27}$$

将(27)代入到(26)中, 便得到关于 z 的方程

$$U(z)-\lambda_0 z+t[V(x_0+z)-f_0(V(x_0+z))(x_0+z)]=\mathbf{0}. \tag{28}$$

这个方程左端的两个加项均属于子空间 $\mathbf{X}_0$, 从而算子 T 可作用于(28); 结果就给出

$$\pi(z)+tR(z)=\mathbf{0}, \tag{29}$$

其中

$$\pi(z)=z, R(z)=T[V(x_0+z)-f_0(V(x_0+z))(x_0+z)].$$

应用定理 2.6 于方程(29). 由于 $z_0=\mathbf{0}$, 故得

$$\pi'(z_0)=I,\quad \Gamma_0=I,\quad \pi''(z)=\mathbf{0},$$

$$R(z_0)=T[V(x_0)-f_0(V(x_0))x_0],$$

$$R'(z)(x)=T[(V(x)-f_0(V(x))(x_0+z)-f_0(V(x_0+z))x],$$

$$R''(z)(x,\tilde{x})=-T[f_0(V(x))\tilde{x}+f_0(V(\tilde{x}))x]$$

$$=-f_0(V(x))T(\tilde{x})-f_0(V(\tilde{x}))T(x).$$

所以可以取

$$\eta = \|T\|\|V\|[1+\|f_0\|\|x_0\|]\|x_0\|,\quad B=1,$$
$$\alpha = \|T\|\|V\|[1+2\|f_0\|\|x_0\|],$$
$$K=0,\quad L=2\|T\|\|V\|\|f_0\|.$$

因此，方程(29)的可解性条件就是

$$h_t = \frac{L\eta|t|^2}{(1-\alpha|t|)^2} \leqslant \frac{1}{2},$$

而这一条件是和(25)等价的.

以 z_t 表示方程(29)的解，则

$$\|z_t\| = \|z_t - z_0\| \leqslant \frac{1-\sqrt{1-2h_t}}{h_t}\eta|t| \leqslant 2\eta|t| = B|t|,.$$

而从等式(27)我们就得到

$$|\lambda_t - \lambda_0| = |\delta| \leqslant |t|\|f_0\|\|V\|\|x_0 + z_t\| \leqslant A|t|.$$

定理完全得证.

注. 我们还要指出一个重要的特殊情况，在这种情况下 η 和 α 可有更为简单的表达式. 设 $\mathbf{X}$ 是 Hilbert 空间而 U 是自共轭算子. 不难验证，这时 f_0 可取由元素 x_0 所定义的泛函

$$f_0(y) = (y, x_0) \qquad (y \in \mathbf{X}).$$

注意，$\|x_0\|^2 = f_0(x_0) = 1$，则

$$\|R(z_0)\| \leqslant \|T\|\|Vx_0 - (Vx_0, x_0)x_0\| \leqslant \|T\|\|V(x_0)\| \leqslant \|T\|\|V\|,$$
$$\begin{aligned}\|R'(z_0)(x)\| &\leqslant \|T\|\|Vx - (Vx, x_0)x_0 - (Vx_0, x_0)x\| \\ &\leqslant \|T\|[\|Vx - (Vx, x_0)x_0\| + \|(Vx_0, x_0)x\|] \\ &\leqslant \|T\|[\|Vx\| + |(Vx_0, x_0)|\|x\|] \\ &\leqslant 2\|T\|\|V\|\|x\|,\end{aligned}$$

因而这时可取

$$\eta = \|T\|\|V\|,\quad \alpha = 2\|T\|\|V\|,$$

而方程(29)的可解性条件就是

$$\frac{2(\|T\|\|V\||t|)^2}{(1-2\|T\|\|V\||t|)^2} \leqslant \frac{1}{2},$$

这个条件就允许我们用不等式

$$\|T\|\|V\||t|\leqslant\frac{1}{4}$$

来代替(25).

М.К.Гаврин[3]直接研究了这种情况并将上述不等式化归成

$$\|T\|\|V\||t|\leqslant\frac{1}{2},$$

而这个条件已经是不可再改善了.

3.7. 考察两个变数的二阶拟线性微分方程

$$A(s,t,u,p,q)\frac{\partial^2 u}{\partial s^2}+B(s,t,u,p,q)\frac{\partial^2 u}{\partial s\partial t}+C(s,t,u,p,q)\frac{\partial^2 u}{\partial t^2}+D(s,t,u,p,q)=0, \tag{30}$$

其中 p 和 q 表示一阶导数 $p=\frac{\partial u}{\partial s}$, $q=\frac{\partial u}{\partial t}$.

设 Q 是一有界区域而 γ 是 Q 的边界. 我们在 Q 上来考察方程(30)并求它在曲线 γ 上满足某些边界条件的解.

如果已知这个问题的近似解 $u_0=u_0(s,t)$,应用Newton法于该问题则可以指出解的存在性条件和解的分布区域(见Кошелев[1],[2]).

我们将假定方程(30)是椭圆型的,确切地讲,将假定对所有的变量,或至少对方程中所遇到的变量来说不等式

$$B^2-4AC<0 \tag{31}$$

满足.

为应用 Newton 法于方程(30),我们将它写成

$$P(u)=\mathbf{0} \tag{32}$$

的形式, 其中 P 是从某个函数空间$\mathbf{U}$映射到另一个函数空间 $\mathbf{V}$ 的

算子. 这两个函数空间将在以后具体化.

设 $F=F(s,t,u,p,q)$ 是某个函数. 我们将用 F_0 表示分别以 u_0, $p_0=\dfrac{\partial u_0}{\partial s}$, $q_0=\dfrac{\partial u_0}{\partial t}$ 代替 F 中的 u,p,q 而得到的结果.

不提出推演的依据问题, 容易得到

$$
\begin{aligned}
P'(u_0)(u)=&A_0\frac{\partial^2 u}{\partial s^2}+\left[\left(\frac{\partial A}{\partial u}\right)_0 u+\left(\frac{\partial A}{\partial p}\right)_0 p+\left(\frac{\partial A}{\partial q}\right)_0 q\right]\frac{\partial^2 u_0}{\partial s^2}\\
&+B_0\frac{\partial^2 u}{\partial s\partial t}+\left[\left(\frac{\partial B}{\partial u}\right)_0 u+\left(\frac{\partial B}{\partial p}\right)_0 p+\left(\frac{\partial B}{\partial q}\right)_0 q\right]\frac{\partial^2 u_0}{\partial s\partial t}\\
&+C_0\frac{\partial^2 u}{\partial t^2}+\left[\left(\frac{\partial C}{\partial u}\right)_0 u+\left(\frac{\partial C}{\partial p}\right)_0 p+\left(\frac{\partial C}{\partial q}\right)_0 q\right]\frac{\partial^2 u_0}{\partial t^2}\\
=&\left[A_0\frac{\partial^2}{\partial s^2}+B_0\frac{\partial^2}{\partial s\partial t}+C_0\frac{\partial^2}{\partial t^2}\right]u+\cdots,
\end{aligned}
\tag{33}
$$

上式中没有写出来的项中不包含关于 u 的二阶导数. 由此可见, $P'(u_0)$ 是线性椭圆型微分算子, 因而对方程(32)应用 Newton 法的每一步都归结为求解线性椭圆型微分方程

$$
P'(u_n)(u_{n+1})=P'(u_n)(u_n)-P(u_n) \qquad (n=0,1,\cdots). \tag{34}
$$

如果利用简化 Newton 过程, 则每一步同样得到椭圆型方程

$$
P'(u_0)(u_{n+1})=P'(u_0)(u_n)-P(u_n) \qquad (n=0,1,\cdots), \tag{35}
$$

不过已经是带有不变的左端了.

同样容易求得(当然, 仍和前面一样不作论证)算子 P 的二阶导数, 这是研究 Newton 法收敛性所必须的. 亦即, 容易看出 $P''(u)(x,y)$ 是 x,y 以及 x 和 y 的直到二阶导数的双线性型, 而且这些导数的系数依赖于 u. 事实上, 包含对 x 的二阶导数与对 y 的二阶导数的乘积的那些加项在所指出的公式中是没有的.

现在来谈谈空间 **U** 和 **V**. 假定这两个空间受下述要求的约束:

1) 算子 P 将空间 $\mathbf{U}$(或它的一部分 Ω) 映射到空间 $\mathbf{V}$.

2) 算子 P 应该二次可微并且 P 的导数由以上所指出的表达式得到.

3) 对无论怎样的 $u\in\Omega$, 双线性型 $P''(u)(x,y)$ 的系数 F_k(看作是乘算子)是从 $\mathbf{V}$ 到 $\mathbf{V}$ 的连续线性算子并且

$$\|F_k\|_{\mathrm{V}}^{\mathrm{V}}\leqslant M \quad (u\in\Omega). \tag{36}$$

类似地, 如果 $x\in\mathbf{U}$, 则函数 $\frac{\partial^{\alpha+\beta}x}{\partial s^\alpha\partial t^\beta}(\alpha,\beta=0,1;\alpha+\beta\leqslant1)$ (同样视作乘算子)是从 $\mathbf{V}$ 到 $\mathbf{V}$ 的连续线性算子并且

$$\left\|\frac{\partial^{\alpha+\beta}x}{\partial s^\alpha\partial t^\beta}\right\|_{\mathrm{V}}^{\mathrm{V}}\leqslant M_1\|x\|_{\mathrm{U}}. \tag{37}$$

4) 将元素 $x\in\mathbf{U}$ 和它的导数 $\frac{\partial^{\alpha+\beta}x}{\partial s^\alpha\partial t^\beta}(\alpha,\beta=0,1,2;\ \alpha+\beta\leqslant2)$ 对应起来的算子是从 $\mathbf{U}$ 到 $\mathbf{V}$ 的连续线性算子.

5) 算子 $P'(u_0)$ 存在连续逆算子 $\Gamma_0=[P'(u_0)]^{-1}$ (从空间 $\mathbf{V}$ 映射到空间 $\mathbf{U}$).

满足上述要求的空间 $\mathbf{U}$ 和 $\mathbf{V}$ 称为对应空间. 只要初始逼近选取得很成功, 则可对这样的空间偶应用定理 1.6. 事实上, 条件 3) 和 4)就保证了 $\|P''(u)\|$ 的一致有界性, 例如, 对于形如 $F\frac{\partial x}{\partial s}\cdot\frac{\partial^2 y}{\partial t^2}$ 的项, 根据(36)和(37)

$$\begin{aligned}\left\|F\frac{\partial x}{\partial s}\frac{\partial^2 y}{\partial t^2}\right\|_{\mathrm{V}}&\leqslant\|F\|_{\mathrm{V}}^{\mathrm{V}}\left\|\frac{\partial x}{\partial s}\right\|_{\mathrm{V}}^{\mathrm{V}}\left\|\frac{\partial^2}{\partial t^2}\right\|\|y\|_{\mathrm{U}}\\&\leqslant M\cdot M_1\left\|\frac{\partial^2}{\partial t^2}\right\|\|x\|\|y\|.\end{aligned}$$

下面我们将指出某些具体的对应空间偶. 不过仅在第一个最简单的情形下来谈到对条件 1)—5)的验证.

考察方程[*)]

$$P(u)\equiv\pi(u)+R(u)\equiv\Delta u+\frac{\partial}{\partial s}(pH)+\frac{\partial}{\partial t}(qH)=\varphi$$

$$\left(\pi(u)=\Delta u,\ R(u)=\frac{\partial}{\partial s}(pH)+\frac{\partial}{\partial t}(qH)\right) \tag{38}$$

并求在正方形 $Q=[0,\pi;0,\pi]$ 的边界上满足条件

$$\frac{\partial u}{\partial n}=0 \tag{39}$$

的 Neumann 问题的解.

关于函数 H 我们将假定它本身以及一定阶的导数连续并且 H 有界.

由 Green 公式可知, 若 u 满足边界条件(39), 则

$$\iint_Q [P(u)]\,ds\,dt=0; \tag{40}$$

因此, 函数 φ 应满足条件 $\iint_Q \varphi\,ds\,dt=0$.

在考察方程(38)的同时, 我们再考察一对空间

$$\mathbf{U}=\overset{\circ}{\mathbf{W}}{}_2^{(3)},\qquad \mathbf{V}=\overset{\circ}{\mathbf{W}}{}_2^{(1)},$$

其中 $\overset{\circ}{\mathbf{W}}{}_2^{(3)}$ 由满足边界条件(39)和条件

$$\iint_Q u(s,t)\,ds\,dt=0 \tag{41}$$

的函数 $u\in\mathbf{W}_2^{(3)}$ 组成. 空间 $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ 由满足条件

$$\iint_Q v(s,t)\,ds\,dt=0 \tag{42}$$

的函数 $v\in\mathbf{W}_2^{(1)}$ 组成. 这两个空间中的范数取为

---

*) 例如, 在薄板问题中遇到的极小曲面方程

$$\frac{\partial}{\partial s}\frac{p}{\sqrt{1+p^2+q^2}}+\frac{\partial}{\partial t}\frac{q}{\sqrt{1+p^2+q^2}}=0$$

就是这种形式的方程.

$$\|u\|=\left[\iint_Q\sum_{\alpha+\beta=3}\left(\frac{\partial^3 u}{\partial s^\alpha\partial t^\beta}\right)^2 ds\,dt\right]^{\frac{1}{2}},$$

$$\|v\|=\left[\iint_Q\left(\left(\frac{\partial v}{\partial s}\right)^2+\left(\frac{\partial v}{\partial t}\right)^2\right)ds\,dt\right]^{\frac{1}{2}}\qquad(u\in\mathbf{U},\ v\in\mathbf{V}).$$

我们来验证空间偶 U, V 满足条件 1)—5) (Ω 取为半径充分大的球).

由嵌入定理可知函数 $u\in\mathbf{U}$ 的二阶导数是任意次幂可积的；因而特别有

$$\left[\iint_Q\left(\frac{\partial^2 u}{\partial s^2}\right)^4 ds\,dt\right]^{\frac{1}{4}}<M_2\|u\|,\tag{43}$$

对于其他二阶导数亦有类似的结果. 函数 u 本身以及它的一阶导数 p 和 q 有界.

我们来考察 $\|P(u)\|$ 的表达式中有特别代表性的一项:

$$\left\{\iint_Q\left[\frac{\partial}{\partial s}\left(H\frac{\partial^2 u}{\partial s^2}\right)\right]^2 ds\,dt\right\}^{\frac{1}{2}}\leqslant\left\{\iint_Q\left(H\frac{\partial^3 u}{\partial s^3}\right)^2 ds\,dt\right\}^{\frac{1}{2}}$$

$$+\left\{\iint_Q\left[\left(\frac{\partial H}{\partial s}+\frac{\partial H}{\partial u}p+\frac{\partial H}{\partial p}\frac{\partial^2 u}{\partial s^2}+\frac{\partial H}{\partial q}\frac{\partial^2 u}{\partial s\partial t}\right)\frac{\partial^2 u}{\partial s^2}\right]^2 ds\,dt\right\}^{\frac{1}{2}}$$

$$<\max|H|\,\|u\|+\max\left|\frac{\partial H}{\partial q}\right|\left\{\iint_Q\left(\frac{\partial^2 u}{\partial s\partial t}\right)^4 ds\,dt\right\}^{\frac{1}{4}}$$

$$\times\left\{\iint_Q\left(\frac{\partial^2 u}{\partial s^2}\right)^4 ds\,dt\right]^{\frac{1}{4}}+\cdots$$

由于(43), 上不等式右端的第二项有界, 没有明确写出来的那些项显然也有界. 从而

$$P(u)\in\mathbf{V}.$$

算子 P 的可微性以及 $P'(u_0)$ 具有如(33)那样的表达式均可基

于同样的理由加以验证．首先规定表达式(33)是弱导数，又因为 P' 连续，则 P 当然是可微的．同样，P 的二阶导数也是可微的．

我们现在来考察函数 F（双线性型 $P''(u)(x, y)$ 中的一个系数）．显然，F 是 H 的偏导数，因而连续有界．对于 $v\in\mathbf{V}$，有

$$
\left\{\iint_Q\left[\frac{\partial}{\partial s}(Fv)\right]^2 ds\,dt\right\}^{\frac{1}{2}}=\left\{\iint_Q\left[F\frac{\partial v}{\partial s}+v\left(\frac{\partial F}{\partial s}+\frac{\partial F}{\partial u}\frac{\partial u}{\partial s}\right.\right.\right.
$$

$$
\left.\left.\left.+\frac{\partial F}{\partial p}\frac{\partial^2 u}{\partial s^2}+\frac{\partial F}{\partial q}\frac{\partial^2 u}{\partial s\partial t}\right)\right]^2 ds\,dt\right\}^{\frac{1}{2}}\leqslant\max|F|\,\|v\|
$$

$$
+\left\{\iint_Q\left[v\left(\frac{\partial F}{\partial s}+\frac{\partial F}{\partial u}\frac{\partial u}{\partial s}+\frac{\partial F}{\partial p}\frac{\partial^2 u}{\partial s^2}+\frac{\partial F}{\partial q}\frac{\partial^2 u}{\partial s\partial t}\right)\right]^2 ds\,dt\right\}^{\frac{1}{2}}. \tag{44}
$$

根据嵌入定理，函数 v 是任意次幂可积的，特别有

$$
\left\{\iint_Q[v(s, t)]^4 ds\,dt\right\}^{\frac{1}{4}}\leqslant M_3\|v\|.
$$

根据这个不等式，我们来估计(44)中第二个加项中的一项．例如，

$$
\left\{\iint_Q\left[v\frac{\partial F}{\partial p}\frac{\partial^2 u}{\partial s^2}\right]^2 ds\,dt\right\}^{\frac{1}{2}}
$$

$$
\leqslant\max\left|\frac{\partial F}{\partial p}\right|\left\{\iint_Q v^4 ds\,dt\right\}^{\frac{1}{4}}\left\{\iint_Q\left(\frac{\partial^2 u}{\partial s^2}\right)^4 ds\,dt\right\}^{\frac{1}{4}}
$$

$$
\leqslant M_4\|v\|.
$$

由此可见 $\mathscr{F}$ 是从 $\mathbf{V}$ 到 $\mathbf{V}$ 的连续线性算子并且满足(36)．条件 3）的第二部分由函数 $u\in\mathbf{U}$ 及其导数 p 和 q 的有界性推知．

条件4）是显然的．

现在来看条件 5）．考虑到要用到 2.6 的结论，因此只须证明 Laplace 算子（看作是从 $\mathbf{U}$ 到 $\mathbf{V}$ 的算子）有连续逆算子存在就可以了．为此，我们来看方程

$$\Delta u = v \qquad (v \in \mathbf{V}). \tag{45}$$

将上面右端的函数 $v(s, t)$ 展成余弦级数：

$$v(s, t) = \sum_{k, m=0}^{\infty} a_{km} \cos ks \cos mt.$$

由于(42)，$a_{00} = 0$. 因此，容易看出满足边界条件(39)的方程(45)的唯一解由级数

$$u(s, t) = \sum_{k, m=0}^{\infty} \frac{a_{km}}{k^2 + m^2} \cos ks \cos mt$$

给出．现在来验证 $u \in \mathbf{U}$. 事实上，由于 v 的导函数平方可积，故有

$$\sum_{k, m=0}^{\infty} (k^2 + m^2)\ a_{km}^2 = M_4 \iint_Q \left[\left(\frac{\partial v}{\partial s}\right)^2 + \left(\frac{\partial v}{\partial t}\right)^2 \right] ds dt$$

$$= M_4 \|v\|^2.$$

但由函数 u 的三阶导数所定义的级数，例如级数

$$\frac{\partial^3 u}{\partial s^2 \partial t} = \sum_{k, m=0}^{\infty} \frac{k^2 m}{k^2 + m^2} a_{km} \cos ks \cos mt,$$

显然是平均收敛的．因此

$$\iint_Q \left(\frac{\partial^3 u}{\partial s^2 \partial t}\right)^2 ds dt = M_5 \sum_{k, m=0}^{\infty} \frac{k^4 m^2}{(k^2 + m^2)^2} a_{km}^2$$

$$\leqslant M_6 \sum_{k, m=0}^{\infty} k^2 a_{km}^2 \leqslant M_7 \|v\|^2.$$

所以 $u \in \mathbf{U}$ 并且 $\|u\| \leqslant M_8 \|v\|$；算子 Δ^{-1} 的存在性得证．

这样，我们就可以应用 Newton 法的收敛性定理．特别要指出的是，如果代替(38)而考察带有参数 μ 的方程

$$\Delta u + \mu \left[\frac{\partial}{\partial s}(pH) + \frac{\partial}{\partial t}(qH) \right] = \varphi,$$

则因 $\mu = 0$ 时解是存在的，所以根据定理 2.6，对于所有充分小的

μ 来说上述方程的解也都是存在的.

应当指出的是，应用定理 2.6 的格式求解上述方程归结为在过程的每一步解 Poisson 方程.

最后，我们再概略地指出两对对应空间. 如果求解 Dirichlet 问题，可取 $\mathbf{U}=\overset{\circ}{\mathbf{W}}{}_p^{(2)}, \mathbf{V}=\mathbf{L}^p$. 这时，条件 1)—4) 的验证并不很难. 但对条件 5)则不然，它的验证相当复杂.

最后，作为 $\mathbf{U}$ 可以取空间 $\mathrm{Lip}^{(2)}\alpha$(由所有连续可微的且其二阶导数满足以 α 为指数的 Lipschitz 条件的函数组成的空间)，相应的$\mathbf{V}$应该取为$\mathrm{Lip}\alpha$. 和前面一样，这里的主要困难在于对条件5)的验证. 要验证它，需要利用偏微分方程论中的一些精细的定理.

Newton 法的其他应用除了以上所述以外还可见Мысовских [2]，及 Николаев[1]的著作.

§ 4. 格-赋范空间中的 Newton 法

在 § 1 中所述的对于赋范空间的 Newton 法形式，可在更广的一类空间中进行，这种空间是与向量格有密切联系的格-赋范空间类. 4.1 和 4.2 的结果属于 Л. В. Канторович[4]，[10]，[12]，4.3 的结果属于 А. И. Балуев [1]. 在叙述这些结果时我们仿照 Вулих-I，在这本书中可以见到这里所述定理的详细证明及所引用的文献.

4.1. 设 $\mathbf{X}$ 是实向量空间，$\mathbf{Z}$ 是向量格，如果对于任一元素 $x\in\mathbf{X}$ 对应于叫做该元素的抽象范数的一个正元素 $|x|\in\mathbf{Z}$，并且满足通常的公理：

1) 如果 $x\neq\mathbf{0}$，则 $|x|>\mathbf{0}$;

2) $|\lambda x|=|\lambda||x|$;

3) $|x_1+x_2|\leqslant|x_1|+|x_2|$,

则称 $\mathbf{X}$ 是使用$\mathbf{Z}$ 的格-赋范空间.

向量格 $\mathbf{Z}$ 叫做用以赋范 (Нормирующое) 的向量格. 赋范空间乃是格-赋范空间的特殊情形，换言之，赋范空间是以实数的 K-空间作为用以赋范的向量格而成的格-赋范空间. 更一般地，任一局部凸空间可以看成是利用 K-空间 $s(T)$的格-赋范空间，其中 T 的势等于确定半模组的势. 最后，如果

在任意的向量格X中取其模为元素的抽象范数，则X成为格-赋范空间，其中用以赋范的向量格是X本身.

设X是利用向量格Z的格-赋范空间，如果在X中 $|x_n - x| \xrightarrow{(o\sigma)} 0$ ($x, x_n \in \mathbf{X}$)，则称序列$\{x_n\}$是$(o\mathbf{Z})$收敛于元素$x \in \mathbf{X}$的，记为 $x_n \xrightarrow{(o\mathbf{Z})} x$. 设序列$\{x_n\} \subset \mathbf{X}$，如果在Z中存在序列$z_m \downarrow 0$，使得对所有的$n > m$，

$$|x_n - x_m| \leqslant z_m,$$

则序列$\{x_n\}$称作 $(o\mathbf{Z})$基本的.

如果任一个$(o\mathbf{Z})$基本序列都是$(o\mathbf{Z})$收敛的，则称空间X是$(o\mathbf{Z})$完备的. Banach空间是$(o\mathbf{R}^1)$完备的；K空间X是$(o\mathbf{X})$完备的.

具有在XI. 1.4中所引进的混合范数的空间$\mathbf{X}[\mathbf{Y}]$是重要的格-赋范空间的例子. 如果用

$$|K|(t) = \|K(\cdot, t)\|_{\mathbf{Y}}$$

引进抽象范数，则空间$\mathbf{X}[\mathbf{Y}]$可看作是利用K-空间X的赋范空间.

请读者自行验证空间$\mathbf{X}[\mathbf{Y}]$是$(o\mathbf{X})$完备的.

我们引进一些在后面研究Newton法时要用到的辅助材料.

其次，在整个4.1和4.2的讨论中总假定X和Y是分别利用K-空间Z和W的格-赋范空间.

设U: X->Y是线性算子，U_0: Z->W是正算子，如果对所有的$x \in \mathbf{X}$

$$|U(x)| \leqslant U_0(|x|),$$

则称U_0是算子U的模控制算子.

我们引进算子导数的概念，它在研究泛函方程时也有用.

设P:X->Y是任意的算子，取指定的元素$x_0 \in \mathbf{X}$并设存在这样的线性算子U: X->Y，使得对于任意的$x \in \mathbf{X}$及任意的数列$t_k \to 0$ ($t_k \neq 0$)有

$$(o\mathbf{W})\text{-}\lim \frac{P(x_0 + t_k x) - P(x_0)}{t_k} = U(x),$$

则称线性算子P在点x_0可导，U是算子P在点x_0处的导数，记为

$$U = P'(x_0).$$

现在来考察相应的积分概念. 设$x_0, x \in \mathbf{X}$，用$[[x_0, x]]$来表示联结点x和y的区间(参见II. 3.1，不要将它和序区间混为一谈).

设函数$F(x)$在区间$[[x_0, x]]$上给定，其值是从X到Y中的线性算子.

设给定区间[0, 1]的分割序列$\{\sigma_n\}$：$0 = t_0 < t_1 < \cdots < t_n = 1$，使得$\lambda =$

$\max(t_{k+1}-t_k)\to 0$, 并指定 $\tau_k\in[t_k, t_{k+1}]$ $(0\leqslant k\leqslant n-1)$. 令 $\xi_k=(1-\tau_k)x_0+\tau_k x$. 如果在 **Y** 中积分和序列

$$\sum_{k=0}^{n-1} F(\xi_k)(t_{k+1}-t_k)$$

存在 $(o\mathbf{W})$-极限, 而且此极限与 $\{\sigma_n\}$ 和 $\{\tau_k\}$ 的选择无关, 则称此极限为函数 F 的积分并记为 $\int_{x_0}^{x} F(x)\,dx$.

这一积分定义特别可应用于算子 $P:\mathbf{X}\to\mathbf{Y}$ 的导数 $P'(x)$ 上. 如果对所有的 $x\in\Omega\subset\mathbf{X}$, $P'(x)$ 存在并且对于所有使得 $x_0+\Delta x\in\Omega$ 的 Δx,

$$P(x_0+\Delta x)-P(x_0)=\int_{x_0}^{x_0+\Delta x} P'(x)\,dx,$$

则称在集合 Ω 上对于算子 P 的 Newton-Leibnitz 公式成立.

我们不再推导 Newton-Leibnitz 公式成立的条件. 请读者从 XVII. 1.7 中的证明出发自己建立充分条件(参见 Вулих-I 第 372 页).

4.2. 现在假设 **X** 是 $(o\mathbf{Z})$ 完备的, 考察方程

$$P(x)=0, \tag{1}$$

其中 P 是从 **X** 到 **Y** 的算子. 指定 $x_0\in\mathbf{X}$, 象在 1.1 中那样, 可用 Newton 法来解方程(1):

$$x'_{n+1}=x'_n-[P'(x'_n)]^{-1}(P(x'_n))$$
$$(n=0, 1, 2, \cdots;\ x'_0=x_0),$$

也可用简化的 Newton 法来解:

$$x_{n+1}=x_n-[P'(x_0)]^{-1}(P(x_n)) \tag{2}$$
$$(n=0, 1, 2, \cdots).$$

我们只简短地考察简化 Newton 法. 令 $\Gamma_0=[P'(x_0)]^{-1}$ (而且我们暂且只要求 Γ_0 是在整个 **Y** 上给定的线性算子). 除了方程(1)外还考察方程

$$Q(z)=0, \tag{3}$$

其中 Q 是从 **Z** 到 **W** 的算子, 方程(3)将起控制方程的作用. 假定在 **Z** 中给定序区间 $[z_0, z']$;

$$D=\{x\in\mathbf{X}: |x-x_0|\leqslant z'-z_0\}.$$

定理 1. 设算子 P 和 Q 满足下列条件:

1) 在集 D 上对于算子 P 的 Newton-Leibnitz 公式成立;

2) 在区间 $[z_0, z']$ 上对算子 Q 的 Newton-Leibnitz 公式成立;

3) 如果$\{z_n\}_{n=1}^{\infty}\subset[z_0, z']$且$z_n \xrightarrow{(o)} z$, 则$Q(z_n)\to Q(z)$;

4) 存在在整个$\mathbf{Y}$上给定的逆算子: $\Gamma_0=[P'(x_0)]^{-1}$和在整个$\mathbf{W}$上给定的$\Delta_0=[Q'(z_0)]^{-1}$, 并且$\Delta_0\in\widetilde{\mathbf{L}}_{n\sigma}(\mathbf{W}, \mathbf{Z})$, 其中$\Delta_0$是$\Gamma_0$的模控制;

5) $|\Gamma_0 P(x_0)|\leqslant-\Delta_0 Q(z_0)$;

6) 如果$x\in D$及$z\in[z_0, z']$使得$|x-x_0|\leqslant z-z_0$, 则算子$I-\Delta_0 Q'(z)$是$I-\Gamma_0 P'(x)$的模控制;

7) 方程(3)在区间$[z_0, z']$上有解$\tilde{z}$.

这时, 方程(1)有满足条件$|\tilde{x}-x_0|\leqslant\tilde{z}-z_0$的解$\tilde{x}$, 初值点为$x_0$的简化Newton过程$(o\mathbf{Z})$收敛于解$\tilde{x}$.

如果在定理1的条件中, 方程(3)在区间$[z_0, z']$中有唯一的解, 则方程(1)在集D中也有唯一的解.

此定理的证明可从Вулих-I找到(定理XII.5.1和XII.5.2).

4.3. 如果泛函方程在序空间中求解, 则自然寻找二个单调序列从不同的方向收敛于解. 这里所述的结果可解释为常微分方程理论中的C. A. Чаплыгин法的抽象推广.

考察方程

$$U(x)=\mathbf{0}, \tag{4}$$

其中U是从K-空间$\mathbf{X}$到向量格$\mathbf{Y}$中的算子.

引理1. 设下列条件成立:

1) 存在$x_0, x_0'\in\mathbf{X}$, 使得

$$x_0\leqslant x_0' \text{ 且 } U(x_0)\leqslant\mathbf{0}\leqslant U(x_0');$$

2) 存在将$\mathbf{X}$映射到整个$\mathbf{Y}$上的线性算子T, 使得若$x_0\leqslant x\leqslant x_0'$, 则

$$U(x_0')+T(x-x_0')\leqslant U(x)\leqslant U(x_0)+T(x-x_0); \tag{5}$$

3) 算子T具有正的逆算子T^{-1}.

这时,

a) 由公式

$$x_1=x_0-T^{-1}U(x_0), \quad x_1'=x_0'-T^{-1}U(x_0')$$

所确定的元素x_1和x_1'满足不等式

$$x_0\leqslant x_1\leqslant x_1'\leqslant x_0', \quad U(x_1)\leqslant\mathbf{0}\leqslant U(x_1'); \tag{6}$$

b) 如果方程(4)在序区间$[x_0, x_0']$上有解$\tilde{x}$, 则$x_1\leqslant\tilde{x}\leqslant x_1'$.

证. 因为$U(x_0)\leqslant\mathbf{0}$和$T^{-1}\geqslant\mathbf{0}$, 则$T^{-1}U(x_0)\leqslant\mathbf{0}$, 由此$x_1\geqslant x_0$. 类似地$x_1'\leqslant x_0'$.

其次,

$$x_0'-x_1=x_0'-x_0+T^{-1}U(x_0)=T^{-1}T(x_0'-x)+T^{-1}U(x_0)$$
$$=T^{-1}[T(x_0'-x)+U(x_0)].$$

由不等于(5)的右端及条件 1)

$$T(x_0'-x_0)+U(x_0)\geqslant U(x_0')\geqslant \mathbf{0}.$$

由此 $x_0'-x_0\geqslant \mathbf{0}$, 因而 $x_0\leqslant x_1\leqslant x_0'$, 并且不等式(5)当 $x=x_1$ 时成立. 由此

$$U(x_1)\leqslant U(x_0)+T(x_1-x_0)=U(x_0)-TT^{-1}U(x_0)=\mathbf{0}$$

及

$$U(x_0')+T(x_1-x_0')\leqslant U(x_1)\leqslant \mathbf{0}. \tag{7}$$

其次

$$x_1'-x_1=x_0'-T^{-1}U(x_0')-x_1=T^{-1}T(x_0'-x_1)-T^{-1}U(x_0')$$
$$=-T^{-1}[U(x_0')+T(x_1-x_0')],$$

由此根据(7)可知 $x_1'-x_1\geqslant \mathbf{0}$, 即 $x_1\leqslant x_1'$.

把 $x=x_1'$ 代入不等式(5)的左端得

$$U(x_1')\geqslant U(x_0')+T(x_1'-x_0')=U(x_0')-TT^{-1}(x_0')=\mathbf{0}.$$

从而证明了不等式(6), 余下要验证命题 b) 成立.

设 $x_0\leqslant\tilde{x}\leqslant x_0'$ 和 $U(\tilde{x})=\mathbf{0}$, 这时, 由(5)

$$\tilde{x}-x_1=\tilde{x}-x_0+T^{-1}U(x_0)=T^{-1}[U(x_0)+T(\tilde{x}-x_0)]$$
$$\geqslant T^{-1}U(\tilde{x})=\mathbf{0},$$

由此 $\tilde{x}\geqslant x_1$. 类似地可证得 $\tilde{x}\leqslant x_1$. 引理证毕.

定理 2. 设下列条件成立:

1) $x_0\leqslant x_0'$, $Ux_0\leqslant\mathbf{0}\leqslant Ux_0'$;

2) 如果 $\{\tilde{x}_n\}_{n=1}^{\infty}\subset[x_0, x_0']$ 及 $\tilde{x}_n\xrightarrow{(o)}x$, 则 $U(\tilde{x}_n)\xrightarrow{(o\sigma)}U(x)$;

3) 在区间 $[x_0, x_0']$ 上对于算子 U 的 Newton-Leibnitz 公式成立;

4) 存在将 $\mathbf{X}$ 映射到整个 $\mathbf{Y}$ 的线性算子 T, 使得对于所有的 $x\in[x_0, x_0']$, 对任一个 $h\in\mathbf{X}_+$, 均有

$$[U'(x)](h)\leqslant T(h);$$

5) 算子 T 有正的 (o)-连续逆算子 T^{-1}.

这时, 方程(4)的属于区间 $[x_0, x_0']$ 中的解的集合非空, 其中有最小解 x^* 和最大解 x^{**}. 并且, 如果元素 x_n 和 x_n' 用下列递推公式确定:

$$x_n=x_{n-1}-T^{-1}U(x_n-1)$$
$$x_n'=x_{n-1}'-T^{-1}U(x_{n-1}')\qquad(n\in\mathbf{N}), \tag{8}$$

则 $x_n \uparrow x^*$ 而 $x'_n \downarrow x^{**}$.

证. 我们指出：算子 U 和 T 在每个区间 $[x_n, x'_n]$ 上满足引理 1 的条件.

我们来证明. 对满足下列关系的任意的 $\bar{x}$, x 和 $\bar{x}'$:

$$x_0 \leqslant \bar{x} \leqslant x \leqslant \bar{x}' \leqslant x'_0$$

具有类似于(5)的不等式

$$U(\bar{x}') + T(x - \bar{x}') \leqslant U(x) \leqslant U(\bar{x}) + T(x - x). \tag{9}$$

事实上，利用条件 3) 及积分的定义可得

$$U(x) - U(\bar{x}) = \int_{\bar{x}}^{x} U'(x)dx \leqslant \int_{\bar{x}}^{x} Tdx = T(x - \bar{x}),$$

$$U(\bar{x}') - U(x) = \int_{x}^{\bar{x}'} U'(x)dx \leqslant \int_{x}^{\bar{x}'} Tdx = T(\bar{x}' - x),$$

从而推出不等式(9).

利用不等式(9)及条件 1)，按照归纳法可知对任意的 x_n 和 x'_n，引理 1 的条件成立，由此

$$x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n \leqslant \cdots \leqslant x'_n \leqslant \cdots \leqslant x'_1 \leqslant x'_0.$$

因 $\mathbf{X}$ 是 K-空间，则存在 $x^* = \sup x_n$, $x^{**} = \inf x'_n$ 并且 $x_n \uparrow x^*$, $x'_n \downarrow x^{**}$. 在(8)中取 (o)-极限并考虑到条件 2) 和 5)，可得

$$T^{-1}U(x^*) = T^{-1}U(x^{**}) = \mathbf{0},$$

由此

$$U(x^*) = U(x^{**}) = \mathbf{0},$$

亦即 x^* 和 x^{**} 是方程 (4) 的解. 此外，根据引理 1，方程 (4) 的任意包含在 $[x_0, x'_0]$ 中的解 $\tilde{x}$ 对所有的 n 均满足 $x_n \leqslant \tilde{x} \leqslant x'_n$，所以 $x^* \leqslant \tilde{x} \leqslant x^{**}$.

定理证毕.

泛函分析及其相邻问题方面的专著

Б а н а х (Banach S.).

Théorie des opérations linéaires, Warsaw, 1932. (Украинский перевод: Курс функционального анализу Київ, 1948).

В а й н и к к о Г. М.

Компактная аппроксимация операторов и приближенное решение уравнений. Тарту, Изд. Тартуского гос. ун-та, 1970.

В л а д и м и р о в Д. А.

Булевы алгебры М., «Наука», 1969.

В у л и х Б. З.

I. Введение в теорию полуупорядоченных пространств. М., физматгиз, 1961.

Г а в у р и н М. К.

Лекции по методам вычислений. М., «Наука», 1971.

Г и л ь б е р т (Hilbert D.).

Grundzüge einer allgemeine Theorie der Integralgleichungen. Leipzig, 1912.

Д е м ь я н о в В. Ф. и М а л о з е м о в В. Н.,

Введение в минимакс. М., «Наука», 1972.

Д е м ь я н о в В. Ф. и Р у б и н о в А. М.

Приближенные методы решения экстремальных задач. Л., Издво ЛГУ, 1968.

З а а н е н (Zaanen A. C.).

I. Linear Analysis, Amsterdam, 1960.

З и н г е р (Singer I.).

Bases in Banach spaces, v. I, Springer-Verlag, Berlin — New York, 1970.

И о ф ф е А. Д. и Т и х о м и р о в В. М.

Теория экстремальных задач. М., «Наука», 1974.

К а н т о р о в и ч Л. В. и К р ы л о в В. И.

Приближенные методы высшего анализа. М.—Л., Физматгиз, 1962.

К а р л и н С.

Математические методы в теории игр, программировании и экономике. М., «Мир», 1964.

К о л л а т ц Л.

Функциональный анализ и вычислительная математика. М., «Мир», 1969.

К р а с н о с е л ь с к и й М. А.

I. Топологические методы в теории нелинейных интегральных уравнений. М., Гостехиздат, 1956.

II. Положительные решения операторных уравнений. М., Физматгиз, 1962.

К р а с н о с е л ь с к и й М. А. и др.

Приближенное решение операторных уравнений. М., «Наука», 1969.

К р а с н о с е л ь с к и й М. А. и др.

Интегральные операторы в пространствах суммируемых функций. М., «Наука», 1966.

К р а с н о с е л ь с к и й М. А. и Р у т и ц к и й Я. Б.

Выпуклые функции и пространства Орлича. М., Физматгиз, 1958.

К р ы л о в Н. М.

Методы приближенного решения задач математической физики. Избранные труды, т. 2, 150—204. Изд. АН УССР, Киев, 1961.

К у т а т е л а д з е С. С., Р у б и н о в А. М.

Двойственность Минковского и ее приложения, Новосибирск, «Наука», 1976.

Л а д ы ж е н с к а я О. А.

Краевые задачи математической физики. М., «Наука», 1973.

Л и н д е н ш т р а у с с и Ц а ф р и р и (Lindenstrauss J., Tzafriri L.).

Classical Banach spaces, Lecture Notes in Math., 338, Springer — Verlag, Berlin — Heidelberg-New York, 1973.

Михлин С. Г.

I. Численная реализация вариационных методов. М., «Наука», 1966.

II. Вариационные методы в математической физике. М., «Наука», 1970.

Натансон И. П.

I. Конструктивная теория функций. М., Гостехиздат, 1949.

Никайдо Х. (Nikaito H.)

Выпуклые структуры и математическая экономика. М., «Мир», 1972.

Никольский Н. К.,

Избранные задачи весовой аппроксимации и спектрального анализа, труды Мат. ин-та им. В. А. Стеклова, 120, «Наука», Л., 1974.

Плеснер А. И.

Спектральная теория линейных операторов. М., «Наука», 1965.

Пшеничный Б. Н.,

Необходимые условия экстремума, М., «Наука», 1969.

Рокафеллар Р.

Выпуклый анализ, М., «Мир», 1973.

Рубинштейн Г. Ш.

Конечномерные модели оптимизации, Новосибирск, «Наука», 1970.

Сегё Г (Szegö G.).

Ортогональные многочлены. М., Физматгиз, 1962.

Смирнов В. И.

Курс высшей математики, т. V. М., Физматгиз, 1959.

Соболев С. Л.

I. Некоторые применения функционального анализа в математической физике. Изд. ЛГУ, 1950.

Филиппов А. Ф. и Рябенький В. С.

Об устойчивости разностных уравнений, М., Гостехиздат, 1956.
Хилле Э. и Филлипс Р. (Hille E. Phillips R.).
Функциональный анализ и полугруппы. М., ИЛ, 1962.

本书所使用的文献

А г а х а н о в С. А. и Н а т а н с о н Г. И.

[1] Функция Лебега сумм Фурье — Якоби. Вестник ЛГУ, № 1 (1968), 11—23.

А к и л о в Г. П.

[1] О распространении линейных операций. ДАН СССР, 57, №7 (1947), 643—646.

[2] Необходимые условия распространимости линейных операций. ДАН СССР 59. №3 (1948), 417—418.

[3] О применении одного метода решения нелинейных дифференциальных уравнений к исследованию систем дифференциальных уравнений. ДАН СССР 68, №4 (1949), 645.

А р о н ш а й н и С м и т (Aronszain N., Smith K.).

[1] Invariant subspaces of completely continuous operators. Ann. Math. 60 (1954), 345—350 (Русский перевод: сб. переводов «Математика», 2, № 1 (1958), 97—102).

Б а л у е в А. Н.

[1] О методе С. А. Чаплыгина. Вестник ЛГУ № 13 (1956), 27—42.

Б а н а х (Banach S.).

[1] Sur les opérations dans les ensembles abstraits et leur applications aux equations integrales. Fund. Math. 3(1922), 133—181.

Б и р м а н М. Ш.

[1] Некоторые оценки для метода наискорейшего спуска. УМН 5, № 3 (1950), 152—155.

[2] О вычислении собственных чисел методом наискорейшего спуска. Зап. Ленинград. горного ин-та, 27, №1 (1952), 209—216.

Б и р м а н М. Ш. и С о л о м я к М. З.

[1] Оценки сингулярных чисел интегральных операторов УМН 32, 1

(1977), 17—84.

Б р а у э р (Brouwer L. E. F.)

[1] Uber Abbindung von Mannigfaltigkeiten. Math. Ann. 71(1911), 97—115.

В е р т г е й м Б. А.

[1] О некоторых методах приближенного решения нелинейных операторных уравнений в пространствах Банаха. *УМН 12, 1(1957)*, 166—169.

В л а д и м и р о в В. С.

[1] Математические задачи односкоростной теории переноса частиц. М., «Наука», 1961.

Г а в у р и н М. К.

[1] Uber die Stiltjessche Integration abstrakten Funktionen. Fund. Math. 27 (1936), 254—268.

[2] Аналитические методы исследования функциональных преобразований. Уч. зап. ЛГУ, сер. матем., 19(1950), 59—154.

[3] Об оценках для собственных чисел и векторов возмущённого оператора. *ДАН СССР 96, №6(1954), 1093—1095.*

[4] Нелинейные функциональные уравнения и непрерывные аналоги итеративных методов. Известия вузов, Матем., № 5, (1958), 18—31.

Г а т о (Gateaux R.)

[1] Sur les fonctionelles continues et les fonctionelles analytiques. C. R. Acad. Sci. (Paris) 157(1913), 325—327.

Г о н ч а р о в В. Л.

[1] Теория интерполирования и приближения функций. Изд. 2-е, Гостехиздат, М.—Л., 1954.

Гротендик (Grothendieck A.).

[2] Produits tensoriels topologiques et espaces nucleaires, Mem. Amer. Math. Soc. 16(1955).

Г ю н т е р Н. М.

[1] Теория потенциала. Гостехиздат, М.-Л., 1952.

Д а у г а в е т И. К.

[1] Применение общей теории приближенных методов к исследованию сходимости метода Галеркина для некоторых граничных задач математической физики. ДАН СССР 98, 6 (1954).

[2] О методе моментов для обыкновенного дифференциального уравнения. Сиб. матем. ж. 6, 1 (1965), 70—85.

Д ж е й м с (James R.).

[1] A non-reflexive Banach spaces isometric with its second conjugate spaces. Proc. Nat. Acad U. S. A. 37, № 3 (1951), 174—177.

Д у б о в и ц к и й А. Я. и М и л ю т и н А. А.

[1] Задачи на экстремум при наличии ограничений, ЖВМ и МФ 5, 3 (1965), 395—453.

[2] Необходимые условия слабого экстремума в задачах оптимальных управления со смешанными ограничениями типа неравенств. ЖВМ и МФ 8, № 4 (1968), 725—779.

Е р м о л ь е в Ю. М. и Ш о р Н. З.

[1] О минимизации недифференцируемых функций, «Кибернетика», 1 (1967), 101, 102.

К а к у т а н и (Kakutani S.)

[1] Some characterizations of Euclidean space. Jap. Jorn. Math. 16, № 2 (1939), 93—98.

[2] A generalization of Brouwer's fixed point theorem. Duke Math. J. 8, № 3 (1941).

К а л а н д и я А. И

[1] Об одном прямом методе решения уравнения теории крыла и его применении в теории упругости. Матем, сб. 42 (1957).

К а н т о р о в и ч Л. В.

[1] Об одном новом методе приближенного решения уравнений в частных производных. ДАН СССР 2, №8—9 (1934), 532—536.

[4] О функциональных уравнениях. Уч. зап. ЛГУ 3 (17) (1937), 24—50.

[7] Об одном эффективном методе решения экстремальных задач для квадратичного функционала. ДАН СССР 48, № 7 (1945), 483—487.

[9] Функциональный анализ и прикладная математика. УМН 3, №6 (1948), 89—185.

[10] Принцип мажорант и метод Ньютона. ДАН СССР 76 (1951), 17—20.

[12] Некоторые дальнейшие применения метода Ньютона. Вестник Ленингр. ун-та, № 7(1957), 68—103.

Канторович Л. В., Акилов Г. П., Рубинштейн Г. Ш.

[1] Экстремальные состояния и экстремальные управления. Вестник ЛГУ, № 7(1967), 30—37.

Канторович Л. В. и Вулих Б. З.

[1] Sur la representation des opérations linéaires. Comp. Math. 5 (1937), 119—165.

Канторович Л. В. и Гавурин М. К.

[1] Применение математических методов в вопросах анализа грузопотоков. Сб. статей «Проблемы повышения эффективности работы транспорта». Изд. АН СССР. 1949, стр. 110—138.

Канторович Л. В. и Рубинштейн Г. Ш.

[1] Об одном функциональном пространстве и некоторых экстремальных задачах. ДАН СССР 115, №6 (1957), 1058—1061.

[2] Об одном пространстве вполне аддитивных функций. Вестник ЛГУ, №7 (1958), 52—59.

Карпиловская Э. Б.

[1] О сходимости интерполяционного метода для обыкновенных дифференциальных уравнений. УМН 8, № 3 (1953), 111—118.

[2] О сходимости метода подобластей для обыкновенных интегро-дифференциальных уравнений. Сиб, матем, ж. 4, 3(1963), 632—640.

Карри (Curry H. B).

[1] The method of steepest descent for nonlinear minimization problem. Quart. Appl. Math. 2, 3(1944), 268—271.
Келдыш М. В.
[1] О методе Б. Г. Галеркина для решения краевых задач. Изв. АН СССР, сер. матем., 6(1942), 309—330.
Коротков В. Б.
[1] Интегральные операторы с ядрами, удовлетворяющими условиям Карлемана и Ахиезера I, Сиб. мат. ж. 12, №5 (1971), 1041—1055.
[2] Интегральные представления линейных операторов. Сиб. мат ж. 15, №3 (1974), 529—545.
Кошелев А. И.
[1] Метод Ньютона и обобщенные решения нелинейных уравнений эллиптического типа. ДАН СССР 91, № 6(1953), 1263—1266.
[2] Существование обобщенного решения упруго-пластической задачи кручения. ДАН СССР 99, № 3(1954), 357—360.
Красносельский М. А
[1] Некоторые задачи нелинейного анализа. УМН 9, № 3(1954), 57—114.
[2] О нескольких новых принципах неподвижной точки. ДАН СССР 208, № 6(1973), 1280—1281.
[3] О квазилинейных операторных уравнениях. ДАН СССР 214, № 4(1974), 761—764.
Красносельский М. А., Крейн С. Г.
[1] Итерационный процесс с минимальными невязками. Матем. сб. 31(1952), 315—334.
Красносельский М. А., Чечик В. А.
[1] Об одной теореме Л. В. Канторовича. Тр. семинара по функц. анализу, вып. 3—4, Ростовск. н/Д. ун-т, ВГУ (1960), 50—53.
Левин В. Л.
[1] К задаче о перемещении массы. ДАН СССР 224, №5 (1975), 1016—1019.

[2] Экстремальные задачи с выпуклыми функционалами, полунепрерывными снизу относительно сходимости по мере. ДАН СССР 224, №6 (1975), 1256—1259.

[3] О теоремах двойственности для задачи Монжа-Канторовича, УМН 32, (1977).

[4] Задача Монжа-Канторовича о перемещении массы. В кн. «Методы функционального анализа в математической экономике», М, «Наука», 1977.

Левин А. Ю., Стрыгин В. В.

[1] О быстроте сходимости метода Ньютона — Канторовича. УМН 17, 3(1962), 185—187.

Лерэ Ж. и Шаудер Ю.

[1] Топология и функциональные уравнения. УМН 1, №3—4(1946), 71—95.

Лозановский Г. Я.

[1] Об изоморфных банаховых структурах. Сиб. мат. ж. 10, №1 (1969), 93—98.

[2] О некоторых банаховых структурах. Сиб. мат. ж. 10, №3(1969), 584—599.

[3] О локализованных функционалах в векторных структурах Сб. «Теория функций, функц. ан. и их прилож.», Харьков, вып. 19 (1974), 66—80.

Лозинский С. М.

[1] Пространства $\widetilde{C}_\omega$ и $\widetilde{C}^*_\omega$ сходимость интерполяционных процессов в них. ДАН СССР 59, №8 (1948), 1389—1392.

[2] Обратные функции, неявные функции и решения уравнений. Вестник ЛГУ, №7 (1957), 131—142.

Ломоносов В. И.

[1] Об инвариантных подпространствах семейства операторов, коммутирующих с вполне непрерывным, функц. анализ и его прил. 7, № 3(1973), 55—56.

Любич Ю. И. и Майстровский Г. Д.

[1] Общая теория релаксационных процессов для выпуклых функционалов. УМН 25, № 1(1970), 57—112.

Мысовских И. П.

[1] О сходимости метода Л. В. Канторовича решения функциональных уравнений и его применения. ДАН СССР 70, № 4 (1950), 565—568.

[2] О граничной задаче для уравнения $\Delta u = k(x, y)u^2$, ДАН СССР 94, № 6(1954), 995—998.

Нейман (Neumann J. von).

[1] Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren. Math. Ann. 102(1929), 49—131.

Николаева Г. А.

[1] О приближенном построении конформного преобразования методом сопряженных тригонометрических рядов. Тр. Матем. ин-та им. В. А. Стеклова 53(1959), 236—266.

Никольский С. М.

[1] Линейные уравнения в линейных нормированных пространствах. Изв. АН СССР, сер. матем. 7, № 3(1943), 147—163.

Пелчинский А. и Фигель Т.

[1] О методе Энфло построения банаховых пространств без свойства аппроксимации. УМН 28, № 6(1973), 95—108.

Поляк Б. Т.

[1] Градиентные методы решения уравнений и неравенств. ЖВМ и МФ 4, 6(1964), 995—1005.

[2] Один общий метод решения экстремальных задач. ДАН СССР 174, № 1(1967), 33—36.

Пугачев Б. П.

[1] О сходимости методов, локально близких к методу Ньютона — Канторовича. Тр. семинара по функц. анализу, вып. 7, Воронеж, изд-во ВГУ (1963), 130—136.

Рисс Ф. (Riesz F.).

[3] Leçons sur les systèmes d'equations linéaires à une infinité

d'inconnues, Paris, 1913.

[4] Sur la décomposition des opérations fonctionnelles. Atti Congresso Bologna 3(1928), 143—148.

[5] Sur quelques notions fondamentales dans la théorie générale des opérations linéaires. Ann. of Math. 41 (1940), 174—206. (Русский перевод — УМН 1, № 2(1946), 147—178).

Самокиш Б. А.

[1] Исследование быстроты сходимости метода наискорейшего спуска. УМН 12, № 1 (1957), 238—240.

[2] Метод наискорейшего спуска в задаче о собственных значениях полуограниченных операторов. Изв. вузов, Матем, № 5(1958), 105—114.

Степанов В. В.

[1] Курс дифференциальных уравнений, изд. 8-е. Физматгиз, М. 1959.

Троицкая Е. А.

[1] О собственных числах и собственных векторах вполне непрерывных операторов. Изв. вузов, Матем., № 3(1961), 148—156.

Фреше (Fréchet M.)

[3] Sur les fonctionelles continues. Ann. Ec. Norm. 27, № 3 (1910), 193—216.

Фридман В. М.

[1] О сходимости методов типа наискорейшего спуска. УМН 17, 3 (1962), 201—204.

Харрик И. Ю.

[1] Об одном аналоге неравенства Маркова. ДАН СССР 106, № 2 (1956), 203—206.

[2] О приближении функций, обращающихся на границе области в нуль вместе с частными производными, функциями особого вида, Сиб. матем. ж. 4, № 2(1963), 408—425.

Хеллингер и Тёплиц (Hellinger E., Toeplitz O.).

[1] Grundlagen für eine Theorie der unendlichen Matrizen. Math.

Ann. 69(1910), 289—330.

[2] Integralgleichungen und Gleichungen mit unendlich vielen Unbekannten. Enciclopädie der matematische Wissenschaften, Bd II, C 13(1927), 1335—1616.

Ш а у д е р (Schauder J.)

[1] Zur Theorie stetiger Abbildungen in Funktionalräumen. Math. Zetschr. 26, № 1(1927), 47—65.

[2] Über lineare vollstetige Operatoren. Studia Math. 2(1930), 183—196.

[3] Über den Zusammenhang zwischen der Eindeutigkeit und Lösbarkeit partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus. Math. Ann. 106(1932), 611—721.

[4] Das Anfangswertproblem einer quasilinearen hyperbolischen Differentialgleichung zweiter Ordnung in beliebiger Anzahl von unabhangigen Veränderlichen. Fund. Math., 24 (1935). 213—246.

Э н ф л о (Enflo P.).

[1] A counterexample to the approximation problem in Banach spaces. Acta Math. 130, № 3—4(1973), 309—317 (Русский перевод: сб. переводов «Математика» 8, № 1(1974)).

术语索引

四画

五画

六画

七画

八画

九画

十画

十一画

十二画

十三画

十四画

十六画

记 号 索 引

[General Information]
书名=泛函分析　下册
作者=(苏)Л.В.Канторович　Г.П.Акилов
页数=334
SS号=10098419
出版日期=1982年8月第1版